

A;Cross-references: EMBL:X81696

A;Residues: 1-549 <AME>

A;Residues: 1-549 <AME>


```

Db      32 KPGSSVKSCKAGGTFSSAISWQAPQGLEMM-GGIIPIFGT-ANYAQKFGQRTVI 89
Qy      79 RADSRSLMDQGNFPIIKKLTIEDSDTYICEVEDQKEEVQLL---VFGILTANSDFHL-- 133
Db      90 TADESTS-----YAYMELSLRSEDTAVYYC-----AKTGILPYSSGYPMNDYYVYG 138
Qy      134 --LOGGSLTTLTSPGSSPS-----VOCR----- 156
Db      139 MDVWGCGTTVTVSSGASAPTLFPLVYSCENSPEDTSSVAVGLAODPLPDSTIFSKYKN 198
Qy      157 -----SPRG--KNIQGGKTLVSQLELD-----SGT--WTCTVLQNKQKVEFKDIDVVL 202
Db      199 NSDISSTRGFPVLRGKVAATSQVLLPSKDVMOQIDENHVCKVQHPNGKKNVPLPIY 258
Qy      203 AF--QKASSIVYKKEGEQVEFPLAFYAEKLTGSGELMWQARASSK---SWITFDLK 257
Db      259 AELPPKVSVPVPRDG---FFGNP-----RSKSKLICQAAGFSPROQVSW----- 301
Qy      258 NKEVSVRVTPDPRKLGKTLPLHLTPQALPQVAGSGLTLALAKTGKLGHEVNLVVM 317
Db      302 -----LREGKV-----GSGVTTDOVQAKES----- 324
Qy      318 RATOLQKLTCEVWGPTSPKMLSLKLENKAVSKREKPVVNLNPEAGMOCILSDSQ 377
Db      325 -----GPTTYKVTSTLTLIKESD-----WL---SQSMFTCRVDHRL 357
Qy      378 VLE-SNIKVLPTWSTPVPCPAPBPSCDKHTHCPELGGSPVFLPPPKKDTLMISRP 436
Db      358 TFOONASSMCVPODDTAI-----RVFALPPS-FASIFLTKST 393
Qy      437 EVTCVVAVDVSHEDPEVFNMYVDGVEVHNAKTKPREEQNSTYRVSVLTVLHODMLNG 496
Db      394 KTLCLVTLDTLTYD-STVISTRONGEAVKTHNTISEHPRATISANGEASICDDMNSGE 452
Qy      497 EYAKVSNKALPAPIEKTISKAGQP-REPQVYTLPPSRDELTKQVSLTCLVKGYPS 554
Db      453 RFTCTVHTDLPSPKQKTSIRPKGVALHRPDVLLPAPRQMLRESATITCLVTVGSPA 512
Qy      555 DIAVESNSGQP--ENNYKTTTPVLVLD--SDGSFPLYSKLTVDKSRMOQGVNVEGCVNHEA 610
Db      513 DVPQWQWQROPPLSPPEKYVLSAPMPPEQAPGRYFAHSILTIVSEENWGTETVYCVVAHEA 572
Qy      611 LHNHYTQKSLSPGLQDDETCAPADGELDGLMTT 646
Db      573 LPRNRTERTVDKS-----TEGEVSADDEGEFENIMAT 603

```

RESULT 52

```

S38864
Ig epsilon chain C region - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001
C/Accession: S38864
R/KIP: B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A/Description: Combination of a defined specificity and desired isotype by cloning of ar
A/Reference number: S38864
A/Accession: S38864
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-546 <KIP>
A/Cross-references: EMBL:Z27397; NID:g416537; PIDN:CA81788.1; PID:g940782
F/353-421/Domain: immunoglobulin C region; immunoglobulin homology

```

```

Query Match      11.0%; Score 374; DB 2; Length 548;
Best Local Similarity 23.3%; Pred. No. 1,2e-14;
Matches 155; Conservative 105; Mismatches 206; Indels 198; Gaps 31;
Qy      30 LGKGGDTVELTCTAS--OKKSIOFHW-----KNSNOIKILGNGSFLTKGSPSKLN--- 77
Db      11 LVKGGSLKLTSCAAGLTFSSYGMWVRQIPDKRLLEWAVATISGGTY--TYPPDSVKGRFT 69

```

```

Qy      78 -RADSRSLMDQGNFPIIKKLTIEDSDTYICEVEDQKEEVQLVFGILTANSDFHLQ 136
Db      70 IERDNAKITLY-----LQMSLSKSEDTAMYCAQGVSTMRFAWM-----G 111
Qy      137 QSLTTLTSPGSSPSVOCRSPRGKNIQGGKTLVSQLELDQSG-----TWCTVLQ 188
Db      112 OGTLTVASAGKTTTPSVPLPAG-----SAAQNSMWTLCGLKGVFPPEVTLWNSGSL- 166
Qy      189 NQKVEFKIDIVVLAFOKASSIVYKKEGEQVEFPLAFYAEKLTGSGEL-----WQME 243
Db      167 -----SSGVH-----TPPAVLOSLYTLSSSVTPSSTWPE 198
Qy      244 RASSKSWITPDLKAKNEVSVKRVTPDPKLOMGKKPLHLTPQALPQVAGSGLTLALBA 303
Db      199 TYTCN--VAHASTKVDKTIVRPD---CGCK-----PCIYV----- 230
Qy      304 KTGKLGHEVNLV-----VRAATQOLKLTCEVWGPTSPKMLSLKLENK----- 348
Db      231 -----EVSVFIFPPPKQDVLTRSTIQ--LYCFYGHILNDVSVSMIMDDREITDTLA 281
Qy      349 -----AKVSKREKPVVNLNPEAGM-----CLSDSQVLLSENIKVLPTWSTVPCP 397
Db      282 QTVLKEBQKLASTGSKNLTTEQOMWSESTFCVTSQGVLYLAHTR-----CP 331
Qy      398 ADEPKSCDKHTHCPELGGSPVFLPPPKKDTLMISRTPEVTCVVVDV--SHEDPEVKFN- 455
Db      332 DHEPR-----GVITYLTPSPVL-DLYQNGAPKLTCLVVDLSEKKNVNTVMQ 377
Qy      456 -----WYVDGVEVHNAKTKPREEQNSTYRVSVLTVLHODMLNGEKYKCVSNK 505
Db      378 EKTISVASAGWY---TKHN-----NATTSITSLIPVAKWIMIGGYQCIVDHP 424
Qy      506 ALPAPIEKTISKAGQPREQVYTLPPSRDELTKQVSLTCLVKGYPDIDAVEMNSNQ 565
Db      425 DEPKPIVRITTPGKRSAPPEVYVPPPEE--SEDKRTITCLQNFPPDISQVWLGDK 483
Qy      566 PENN--YKTTTPVLDSGDS---FFLYSKLTVDKSRMOQGVNVEGCVNHEALHN-HYTOKS 619
Db      484 LISNSQHSITTP-LKNSGNSRGFFITSRLEVAKTLMTOKQKFTQVYIHDLQPRLEKT 542
Qy      620 LSLS 623
Db      543 ISTS 546

```

RESULT 53

```

EHMS
Ig epsilon chain C region (version 1) - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
C/Accession: A02144
R/LIU, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.
Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982
A/Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.
A/Reference number: A02144; MUID:83117774; PMID:6818553
A/Accession: A02144
A/Molecule type: mRNA
A/Residues: 1-388 <LIU>
A/Cross-references: GB:000476; NID:g194875; PIDN:AAA8085.1; PID:g387720
C/Complex: An immunoglobulin heterocetramer subunit consists of two identical light (ka
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: duplication; glycoprotein; heterocetramer; immunoglobulin
F/1-44/Domain: immunoglobulin homology (fragment) <IML>
F/81-149/Domain: immunoglobulin homology <IM2>
F/186-254/Domain: immunoglobulin homology <IM3>
F/290-361/Domain: immunoglobulin homology <IM4>
F/10,51,62,133,205,228,332,382/Binding site: carbohydrate (asn) (covalent) #status pred

```

```

Query Match      10.8%; Score 367.5; DB 1; Length 388;
Best Local Similarity 26.7%; Pred. No. 1,9e-14;
Matches 118; Conservative 73; Mismatches 138; Indels 113; Gaps 20;

```



```

Oy 222 SFLAFTVEKLTGSGELMWQAEERASSKSWITFDLKNKEVSKRVTDOPKLOMGKLT--- 278
Db 15 NFPALGSELKVTTSQVTSW-----GKSANQFTC-----HYTHPPSNESRTILVR 59
Oy 279 PLHLTLPALPOYAGSGLTLALEAKTKLHOENVLVVRATLOLOKLTCEVWGPSTSKL 338
Db 60 PNVITEP-----LTFL-----LHSSCPNPFHST---IQLYCFYGIHINDV 98
Oy 339 MLSLKLEKE-----AKYSKREKPVWVNLPEAGMWQ-----CLLSDGQVLL 381
Db 99 SVSNLMMDRELITDITLAQVTLKEBSGLASTGSKLTTEQOMWSESTFTCKVTSQGVLYLA 158
Oy 382 SNIKVLPWSTPVPCEPAPEPKSCDKTHTCPPELLGSPVFLPPPKKDTLMSRTPEVTCV 441
Db 159 HTRR-----CPDHEPR-----GVITYLLPSPPLD-LYQNGAPKLTCL 194
Oy 442 VVDV-SHEDPEVKR-----WYDGEVHNATKPREQYNSTYRVSVLYVLA 489
Db 195 VVDESEKGVVVTWNOEKRTSVSASQWY---TKHN-----NATTSITSLPVVA 241
Oy 490 QDWLNGEKYKGVSKALPAPLIEKISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLYK 549
Db 242 KQWIEGIGQCIIVDPDPKPIVSTKTPGQSAPEVYVPPPEB-SEDKRTLTCLIQ 300
Oy 550 GFVPSDIAVWESNGQPENN--YKTPPVLDSDG--FFLYSKLTVDKSRMQQGVNSFC 604
Db 301 NFPEDISVQWLGSGKLSNGOHSHTTP-LKSNNGSNGFFLFRNLVAKTILMTGRKOPTC 359
Oy 605 SVMHEALHN-HYTKSLSLSPG 625
Db 360 QVHEALQKPRKLEKTIISTSLG 381

```

RESULT 54

```

Ehrt
Ig epsilon chain C region - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1992 #sequence revision 17-Dec-1992 #text_change 16-Jul-1999
C:Accession: A93442; A90937; X02143
R:Hellman, L.; Petersen, U.; Engstrom, A.; Karlsson, T.; Bennich, H.
Nucleic Acids Res. 10, 6041-6049, 1982
A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
A:Reference number: A93442; MUID:83064537; PMID:6292865
A:Accession: A93442
A:Molecule type: mRNA
A:Residues: 1-429 <HEL>
A:Experimental source: strain LOU/C/Wel, immunocytoma IR2
R:Kindvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
DNA 1, 335-343, 1982
A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction,
A:Reference number: A90937; MUID:83182019; PMID:6820340
A:Contents: myeloma IR162
A:Accession: A90937
A:Molecule type: mRNA
A:Residues: 'N', 169-307, 'L', 309-342 <KIN>
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kap)
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1a
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
F:19-80/Domain: immunoglobulin homology <IM1>
F:118-186/Domain: immunoglobulin homology <IM2>
F:223-291/Domain: immunoglobulin homology <IM3>
F:327-398/Domain: immunoglobulin homology <IM4>
F:46,99,170,240,265,369,419/Binding site: carbohydrate (Aan) (covalent) #status predicted

```

```

Query Match 10.6%; Score 363.5; DB 1; Length 429;
Best Local Similarity 30.6%; Pred. No. 3.8e-14;
Matches 91; Conservative 60; Mismatches 99; Indels 47; Gaps 12;
Oy 342 LKLENKAKYSKREKPVWVNLPEAGMWQ-----CLLSDGQVLLSNIKVLPWSTPVP 395
Db 156 IKEKGKLAIVSYR-----LNIIOQWMSSESTFTCKVTSQGE-----NYMAHTRR 199

```

```

Oy 396 CPAPPEKSCDKTHTCPPELLGSPVFLPPPKKDTLMSRTPEVTCVVVDVSHEDPEVKFN 455
Db 200 CSDEPR-----GVITYLLPSPPLD-LYENGPRKLTCLVLDSESE-NITVT 244
Oy 456 WYDGEVHNATKPREQYNSTYRVSVLYTVHQDLNGEKYKGVSKALPAPIEKTI 515
Db 245 WVERKKISIGSASQSRTRKHNAATTSITSLIPVDADWIEGGYOCYRDVHPFPKPIVRSI 304
Oy 516 SKAKGQPREPOVY-LPPSRDELTKNOVSLTCLVKGFPSPDIAYEW--ESNGQPENNYKT 572
Db 305 TKAGKASAPEVYVFLPPEBE--KDKRTLTCLIQNFPEDISVQWLODSGLTKPSQHS 362
Oy 573 TPVLDSDG--FFLYSKLTVDKSRMQGVNSCSVMHEALHN-HYTKSLSLSPG 625
Db 363 TTP-LKXNGSNGRFFISRLVETALMTQKQFICRVIHEALRPRLERTISLSLG 418

```

RESULT 55

```

Ig mu chain - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S25705
R:Patil, S.; Nau, F.
Mol. Immunol. 29, 829-836, 1992
A:Title: Isolation and sequence of a cDNA coding for the immunoglobulin mu chain of the
A:Reference number: S25705; MUID:92342148; PMID:1655560
A:Accession: S25705
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-592 <PAT>
A:Cross-references: EMBL:X59994; NID:G1269; PIDD:CA42611.1; PIDD:G1270
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:483-554/Domain: immunoglobulin homology <IM>

```

```

Query Match 10.6%; Score 363.5; DB 2; Length 592;
Best Local Similarity 22.6%; Pred. No. 5.7e-14;
Matches 155; Conservative 97; Mismatches 23; Indels 181; Gaps 29;
Oy 6 PFRHLVLVQL---ALLPAATQGNKVVLAGKGVLTCTASOKSIQFH--WKNNOIK 60
Db 3 PLMTLLFVLAPRGVLSQVQLQSGPSELVPSFELSLCTVSGSGLTVNHNWIRQASGK 62
Oy 61 ILKNGSFLTKGSKLNDRAISRSL--WDQGNFPLITKMLKEDSDTYICEVEDQKE 117
Db 63 MPWLGVEKGVGNTYVPALPKSRISIRDTSKQVSLSLSMIDDAVYVC----- 114
Oy 118 VQLLVFLGLTNSDTHLIQ-----GQSLTFLT---ESPPGSSPSVOC-RSPRGKNIQ--- 164
Db 115 -----ARSAGVFLADVVDWKGGLLVTVSSBSSEHPKVPFLVCSVSSPSIDENTYALG 166
Oy 165 -----GGKTLVSQTL---ELQDSGTWCT---VLQNKQVVEFKIDIV 200
Db 167 CLARDPVPNVSPSKKNNSTVSSERFWTPPEVLRLDLSWASSQVLAHSSSTFGTGGYL 226
Oy 201 VLAFO--KASSIYKKEGEVSPFLAFT---VEKLTSGELIMQAEASSSKSWITP 254
Db 227 VCEQVHKGEDVGHKGPRREVEVLSPVSVFVPCNSLSGNG-----NSKSLICQAT 279
Oy 255 DLKMKVEVY---KRVTDPLQMGKULPLHLTLPOALPOYAGSGLTLALEAKTGKL 308
Db 280 DFSEKQISLSWDFDGRKIVSD-----ISBQVETVQSSPTTYR 317
Oy 309 HOEVNLVWRATLOKNTLCEVWGPTSPKMLSLKLENKAKYSKREKPVWVNLPEAGMW 368
Db 318 AVSLVLTIREHMSQSAYTQV-----ENKKT----- 345
Oy 369 QCLLSDGQVLLSNIKVLPWSTPVPCEPAPEPKSCDKTHTCPPELLGSPVFLPPPKPK 428
Db 346 -----FGKNAS-----SSCDATPPSP--IG--VFTIPPSFAD 373
Oy 429 TLMISRTPEVTCVVVDVSHEDPEVKFNW-VVDG--VEVHNATKPREQYNSTYRVSVLY 485

```

```

Db      374 -ITLTSATKSLCVTLVLAASD-GLNTSMKSNQKALEFTHY-----FERHLNDFSKARGA 427
      486 TVLHODMLNGEKYCKVSNKALPAPIEKTISKAKG-OPREPOVYTLPPSRDELTKNOVS 543
      428 SVCSSEDMESGEETCTVAHDLTFPEKSAISKPKDVAMKPPSYVYLPTRREQSLTESAS 487
Qy      544 LITLVYGFYPSDIAVEMESNGP--BNNYKTPPVLD--SDGSFELYSKLTYDKSRMOG 599
      488 VTCLVGFAPADVFVQMLQGEVNAKSVTSSPAEDPPSAVFVHSILITLEDMSKG 547
Db      600 NVFSCSVMEALHNHYTKSLSLSPG 625
      548 EYTCVGVGHEALPHMTERTKVDKSTG 573

```

RESULT 56

S03186
Ig heavy chain C region - African clawed frog (fragment)
C/Species: Xenopus laevis (African clawed frog)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C/Accession: S03186
R/Hair: R.N.; Shamblott, M.J.; Amemiya, C.T.; Lilman, G.W.
Nucleic Acids Res. 17, 1776, 1989
A/Title: A second Xenopus immunoglobulin heavy chain constant region isotype gene.
A/Reference number: S03186
A/Accession: S03186
A/Molecule type: mRNA
A/Residues: 1-448 <HA1>
A/Cross-references: EMBL:X13779; NID:G64827; PID:G64828
A/Note: the authors translated the codon TTT for residue 9 as Ser and ATG for residue 34
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: heterotetramer; immunoglobulin

Query Match 10.6%; Score 360.5; DB 2; Length 448;
Best Local Similarity 28.1%; Pred. No. 6e-14;
Matches 110; Conservative 64; Mismatches 141; Indels 77; Gaps 15;

```

Qy      291 YAGSGNLT-LALEAKGKGLHGVNLVYMRATQLOKLTCEVWGPTSPKMLSLKENKA 349
      65 YTLSSQLTLIASWKKSTYKCK--VYHNTNTKOEKSLKLVPCMAPIVQLFQSPQMSD 121
Db      350 KVKSR-----KPVVTLN-----PEAGMQLCLSDSGVILE 381
      122 AISRAGHENINATLDLCTINNRYHQIKYKMLVNGQDVASASVTPSPKEDGYTSSS 181
Qy      382 SNIKVLP-TWSTVPCEPAPKSCDKTH-----CPDLGGPSVFLPFPKPK- 427
      182 SQRLIKGMNKKP-----QYSGIVHTSSNTTIANISQCTEQC-HDMLGYVPLPTPF 234
Db      428 DTLMISTPEYTCV-----VDVSHEDPEV-KFWYVYDVGVANAKTKPREBOYNT 478
      235 HDYFSSNAKYVTCIVSSMKTIENFDISWEEKAGNLEFVEDPLHD-----NGT 284
Qy      479 YRVSVVLTVAHODMLNGEKYCKVSNKALPAPIEKTISKAKG-OPREPOVYTLPPSRDEL 537
      285 YSVASIIISVCAEDMESDCKSVYRSODLPSPVKKTIFFKQNEGTPKAPDVYLLPPSAQEL 344
Db      538 TKNQ-VSLTCLVYKGFYPSDIAVEMESNG--OPENNYKTPPVLDSDGSFELYSKLTVDK 593
      345 IQGEMVLTLCFVVGFGNFKEIFIQWMOGVYSISDEKFIINTVPMKSDSGQYFIYSKALIPA 404
Qy      594 SRMOQNVFSCSVMEALHNHYTKSLSLSPG 625
      405 AKWMOGDVFTCVGVGHEALPLVITQOISDKSSG 436

```

RESULT 57

S00390
Ig gamma chain (clone 36) - chicken (fragment)
N/Alternate names: Ig nu chain
C/Species: Gallus gallus (chicken)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000

C/Accession: S00390
R/Parvari, R.; Aviri, A.; Lentner, F.; Ziv, E.; Tel-Or, S.; Burshtein, Y.; Schechter, EMO J. 7, 739-744, 1988
A/Title: Chicken immunoglobulin gamma-heavy chains: limited VH gene repertoire, combination of JH and JH2 segments
A/Reference number: S00390; MUID:88283642; PMID:3135182
A/Accession: S00390
A/Molecule type: mRNA
A/Residues: 1-504 <PAR>
A/Cross-references: EMBL:X07174
A/Note: this sequence was determined from the differentiated gene
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin

Query Match 10.5%; Score 360; DB 2; Length 504;
Best Local Similarity 23.7%; Pred. No. 7.5e-14;
Matches 147; Conservative 89; Mismatches 212; Indels 172; Gaps 27;

```

Qy      59 IKLNGQSFLLTKGSKLNDRADSRRLMDQGNFPLIKNLIEDSDTYIC----- 109
      3 VAAISSTSGTNYG-SAVKGRATISRDN-CQSTLRQLNLAADDTGYTCARDLGIDL 60
Db      110 ---EVEDQKEVQLVFGITANSDFHLQGSGLTTLSEPPSSSPSSVQCRSPRKNIQGG 166
      61 YAGQIDAMGHGTEVIV--SSASPT-----SPPLVPLSACCSNAPPAVG 104
Qy      167 KTLVSQ---LELQSGTWTCIVLQNKVKERKIDIVLAFKASIVYKKGGEYEFSEF 223
      105 CLLSPPSSAGGIVSESG-----GTVAGKVSCTPVKLSF 138
Qy      224 PLAFVTEKLTSGELMWQAEARSSSKSWTF-----DLNKEVSVKRVTPDPLQMGKTL 278
      139 VRLSPK-----KSFYCSAAPGALLAKKEVQVGRV--D----- 172
Db      279 PLHLTPQALPOYAGSGNLT-LALEAKGKGLHGVNLVYMRATQLOKLTCEVWG--PTSP 336
      173 ---VPPVAPRVQ---VLHASSCTPQSSSEVEL-----CLVTGFSAPASA 210
Qy      337 KMLSLSLKENKAKYKSKREKPVVNLPEAGMQLCLSDSGQV-----LLESNIKVLPT- 389
      211 EV-----EMLVDGVGL--LVASQSPARSGSTYSLSKVNVSCTD 249
Qy      390 ---WSTVPCEPAPKSCDKTHCPDLGGPS-----VFLFPKPKDTLMISRTPEV 438
      250 WRGKSYSGRVNHPNTVYEDHVGCPD--GAGSCSPIQVLAIPSPDE-LTISLDATL 306
Qy      439 TCVVDVSHEDPEVKFWYVYDVGVANAKTKPR--BEQYNSTYRVSVLTVLHODMLNG 496
      307 RCLVNLPL-SDSLSVTWTR--KSGNLRPDMVLQEHFNQGYASASAVPVSTQDWLSGR 363
Qy      497 EYKCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLCLVYGFYPSDI 556
      364 RFTCTVQHEELPLPSKSVRYNTGPTPLIPPAHPELSLSRVTLSCLVAGFEPRI 423
Db      557 AVEW--ESNGOPENNYKTPPVLD-----SDG-SFELYSKLTYDKSRMOQNVFSCG 605
      424 EIRMLRDHRAVATEVTVTLVLPBEERTANGAGGDDGTFVYSGMSVETAKKMGVTFACM 483
Qy      606 VMEALHNHYTKSLSLSPG 625
      484 AVHEALPMRFQSRTLQKQAG 503

```

RESULT 58

MHMSM
Ig mu chain C region, membrane-bound form - mouse
C/Species: Mus musculus (house mouse)
C/Date: 31-Oct-1980 #sequence_revision 30-Jun-1991 #text_change 22-Jun-1999
C/Accession: A02167; A37517; B02166
R/Barly, P.; Rogers, U.; Davis, M.; Calame, K.; Bond, M.; Wall, R.; Hood, L.
Cell 20, 313-319, 1980
A/Title: Two mRNAs can be produced from a single immunoglobulin mu gene by alternative splicing
A/Reference number: A02167; MUID:80222874; PMID:6771020
A/Accession: A02167

A:Molecule type: DNA
 A:Residues: 436-476 <EAS>
 A:Cross-references: GB:V00816; GB:U04444; NID:G52343; PIDN:CAA24197.1; PID:G52344
 R:Rogers, J.; Early, P.; Carter, C.; Calame, K.; Bond, M.; Hood, L.; Wall, R.
 Cell 20, 303-312, 1980
 A:Title: Two mRNAs with different 3' ends encode membrane-bound and secreted forms of in
 A:Reference number: A37517; MUID:8022873; PMID:6771010
 A:Contents: MOPC 104B
 A:Accession: A37517
 A:Molecule type: mRNA
 A:Residues: 410-476 <ROG>
 A:Cross-references: GB:V00821; NID:G52355; PIDN:CAA24202.1; PID:G817972
 R:Kawakami, T.; Takahashi, N.; Honjo, T.
 Nucleic Acids Res. 8, 3933-3945, 1980
 A:Title: Complete nucleotide sequence of mouse immunoglobulin mu gene and comparison wit
 A:Reference number: A02166; MUID:81076590; PMID:6255422
 A:Accession: B02166
 A:Molecule type: DNA
 A:Residues: 1-435, 'GKPTLVNLSIMSDGTCY' <KAN>
 A:Comment: The sequence of residues 1-409 was assumed to be identical with the correspo
 C:Comment: During differentiation, B lymphocytes switch from expression of membrane-bou
 B:
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 14
 C:Superfamily: Immunoglobulin C region; immunoglobulin homology
 C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob
 F:21-91/Domain: immunoglobulin homology <IMM1>
 F:129-201/Domain: immunoglobulin homology <IMM2>
 F:229-307/Domain: immunoglobulin homology <IMM3>
 F:346-417/Domain: immunoglobulin homology <IMM4>
 F:436-476/Domain: carboxyl-terminal <CTS>
 F:14/Disulfide bonds: interchain (to light chain) #status predicted
 F:28-89, 136-199, 246-305, 353-415/Disulfide bonds: #status predicted
 F:46-211, 243, 258, 281/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:216/Disulfide bonds: interchain (to heavy chain) #status predicted
 F:293/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 10.5%; Score 359.5; DB 1; Length 476;
 Best Local Similarity 22.7%; Pred. No. 7.4e-14;
 Matches 127; Conservative 80; Mismatches 187; Indels 165; Gaps 19;

Dy 144 ESPGSSPSVQCRSP-RGKNIQGGKTLVSQLELDQSGTWTCTVLQNKVFKIDIVL 202
 3 QSPFNVPFLVCSBPPLSDKLVAMGCLARDFLPSTISPTN----- 43
 Dy 203 AFQKASSTVYKKEGEQVEFFPLAFYVEKLTGSGELMWAERASSSKSWTFDLKNEVS 262
 44 -----YQNTTEVIQ-----GIRTFPLRTGCKYLA 68
 Dy 263 VKRVTQDPKLOM-----GKKLPLHLTLPLQALPQVAGSGNLTALAKTKGL 308
 69 TSQVLBPKSLIBESDELVLCKHYGGKRNRLHVPITP-----AVAEW 110
 Dy 309 HQEVNLVV-----MRAIOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPVWLN 362
 111 NPNNVNFVPRPDGSGPARPKSLICBATNTTPKPIYWSMKDKLVESGFTTDPVILEN 170
 Dy 363 -----PEA-----GMQCLISDSGQVLBSNIVLPTWSTPVPAPAP 401
 171 KGSTPQRYKVIISTLISEIDWLNINLVYTCRVHRGLFTLK----- 210
 Dy 402 KSCDKTHTCPRLGGPS-----VFLPRPKDITLMISTPRTVCVNVDSHEDPEVKNWY 457
 211 ---NVSTC---AASPSTDIITFTTPSPFAD-IFLSKSAINTCLVSNLATE-TLNLSWA 262
 Dy 458 VDGVEVNAKTPREEQYSTYRVSVLTVLHODWLNGKEKCKVSKKALPAPIETKISK 517
 263 SQSEPLETKIKIMESHPNGTFSKAGVASVCEVDMNRKREVCYVTHRDLPSPKRTISK 322
 Dy 518 AKGQPRE-----POVYTLPSRDEL-T-KNOVSLTCLVKGFPYSDIAVWESNGQ--DENN 569
 323 -----PNEVHKHPRPVYLLPAREQDNLNRESATVTCVVGESPADISQVMIQGLLPQEK 378

Dy 570 YKTPPVLD--SDGSFFLYSKLTVDKSRMOQGNVSCSVMEALAHNYTKSLSPGLQ 627
 379 YVTSAPNPEPAPPEFYTHSLTITEEMNSGERTVCVGHFALPHVTERTVDS----- 434
 Dy 628 LDFTCAPADGDELGLMTT 646
 435 -TEGEVNAEPEGFENLMTT 452
 Db

RESULT 59
 S15590
 Ig heavy chain - human
 C:Species: Homo sapiens (man)
 C:Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 C:Accession: S15590
 R:Neale, G.A.M.; Kitchingman, G.R.
 Nucleic Acids Res. 19, 2427-2433, 1991
 A:Title: mRNA transcripts initiating within the human immunoglobulin mu heavy chain enh
 A:Reference number: S15590; MUID:9125286; PMID:1904154
 A:Accession: S15590
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-474 <NEA>
 A:Cross-references: EMBL:X58529
 A:Note: the authors translated the codon CAA for residue 265 as Glu
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:344-415/Domain: immunoglobulin homology <IMM>

Query Match 10.5%; Score 359; DB 2; Length 474;
 Best Local Similarity 26.7%; Pred. No. 7.9e-14;
 Matches 86; Conservative 65; Mismatches 121; Indels 50; Gaps 11;

Dy 332 GPISPKMLSKLENKAKVSKREKPVVNLNPEAKMQCLISDSGQVLE-SNIKVLPTW 390
 172 GPITYKVTSLTIKESD-----WL---SQSMFTCRVDRHGLTFQGNASMCVPDQ 218
 Db
 Dy 391 STVPKCAPPKSKCDKHTHCPBELLGSPSVLFPKPKDITLMISTPRTVCVNVDSHEDP 450
 219 DTAI-----RVFAIPPS-FASIFLTSTKLTCLVTLDTTYD- 253
 Dy 451 EVKENWYVDGVEVNAKTPREEQYSTYRVSVLTVLHODWLNGKEKCKVSKKALPAP 510
 254 SVTISWRQNGQAVKHTNISESHNATPFAVGAASICEEDMNSGERTVCVTHRDLPSP 313
 Db
 Dy 511 IETKISAKQOP-REPQVYTLPSRDEL-T-KNOVSLTCLVKGFPYSDIAVWESNGQ-- 566
 314 LKQITSPKGVALHPRPVYLLPAREQDNLNRESATITCLVTGSPADVFMQMRGQPLS 373
 Dy 567 ENNYKTPPVLD--SDGSFFLYSKLTVDKSRMOQGNVSCSVMEALAHNYTKSLSP 624
 374 PEKYTSAPNPEPAPPEFYTHSLTITEEMNSGERTVCVGHFALPHVTERTVDS- 432
 Dy 625 GLQDFTCAADGDELGLMTT 646
 433 -----TEGEVNAEPEGFENLMTT 450
 Db

RESULT 60
 S37768
 Ig mu chain C region - human
 C:Species: Homo sapiens (man)
 C:Date: 12-Feb-1998 #sequence_revision 12-Feb-1998 #text_change 23-Jul-1999
 C:Accession: S37768
 R:Harindranath, N.; Donadel, G.; Sigounas, G.; Nockins, A.L.
 Mol. Immunol. 30, 111-112, 1993
 A:Title: Comparison of complete nucleotide sequence of the human IgM heavy chain const
 A:Reference number: S37767; MUID:93109369; PMID:8417370
 A:Accession: S37768
 A:Molecule type: mRNA
 A:Residues: 1-453 <HAR>
 A:Cross-references: EMBL:X67301; NID:G38407; PIDN:CAA47714.1; PID:G38408
 A:Experimental source: cell line Ab 63

C:Genetics:
 A:Map position: 14q32
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: glycoprotein; heterodimer; immunoglobulin
 C:Date: 31-Mar-1998 #sequence_revision 13-Jun-1998 #text_change 16-Jul-1999
 F:21-90/Domain: immunoglobulin homology <IM1>
 F:127-199/Domain: immunoglobulin homology <IM2>
 F:237-305/Domain: immunoglobulin homology <IM3>
 F:344-415/Domain: immunoglobulin homology <IM4>
 F:14/Dsulfide bonds: interchain (to light chain) #status predicted
 F:28-88,134-197,244-303,351-413/Dsulfide bonds: #status predicted
 F:16,209,272,279,440/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:214,452/Dsulfide bonds: interchain (to heavy chain) #status predicted
 F:291/Dsulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 10.5%; Score 357.5; DB 2; Length 453;
 Best Local Similarity 27.6%; Pred. No. 9.1e-14;
 Matches 83; Conservative 62; Mismatches 11; Indels 45; Gaps 10;

332 GPTSPKLMSTIKENKAKVSKREKRVWVLPNPAAGMOCCLSDSGQVLE-SNIXVLPW 390
 172 GPTTVVSTLTIKESD-----WL---SQSMFTCKVDHRLGTFQNMSSMCVPO 218
 391 STEVPCAPPEPKSCDKHTCTCELLGSPSYFLPPPKEDTLMSRTPEVTCVVDVSHEDP 450
 219 DPAI-----RVPALPPS-FASIFLTKSKTKLTCLVTLDTTD- 253
 451 EVKFNMYVDGVHNAKTKPREQVNSTRVSVLTVLHODMLNGEKYCKKSNKALPAP 510
 254 SVTISWTRONGEAVKHTNISSEHPNATFSAVEASICEBDNMSGERTCTVHTDLPSP 313
 511 IETIKKAKGP-REPOVYTLPPSRDELTKNOVSLTCLVGYPSYDIAVEMESGQP-- 566
 314 LKQTFIRPKVLAHRPVDVLLPAREQMLNRESATITCLVTGSPADVOMORQDPDS 373
 567 ENNYKTPPVLD--SDGSFELYSKLTIVDKSRWQGNVFCSCVHGAHLNHYTQKSLSP 624
 374 PEKYVTSAPPEPQAGRYFAHSILTVSEEMTGTGYTCVVAHEALPKRVTERYDCKST 433

QY 625 G 625
 DB 434 G 434

RESULT 61
 I60082
 CD4 receptor - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 23-Jul-1999
 C:Accession: I60082
 R:Zverev, V.V.; Sidarov, A.V.; Nedospasov, S.A.; Malushova, V.V.; Udalova, I.A.; Andzha
 Vopr. Virusol. 40, 100-102, 1995
 A:Title: [Nucleotide sequence of two exons of the human T-lymphocyte CD4 receptor gene].
 A:Reference number: I60082; MUID:95407135; PMID:767667
 A:Accession: I60082
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-71 <RES>
 A:Cross-references: GB:S79267; NID:91086922; PIDN:AAB35273.1; PID:91086923
 C:Genetics:
 A:Insertions: 17/1
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

Query Match 10.5%; Score 357; DB 2; Length 71;
 Best Local Similarity 98.6%; Pred. No. 9.2e-15;
 Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLTQALTPAATGKNKLVGKGGDTVELCTASOKKSIQFHKNSNQK 60
 DB 1 MNRGVPFRHLVLTQALTPAATGKNKLVGKGGDTVELCTASOKKSIQFHKNSNQK 60

QY 61 ILGNOSFLTK 71
 DB 61 ILGNOSFLTK 71

RESULT 62
 EHNU
 Ig epsilon chain C region - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Mar-1998 #sequence_revision 13-Jun-1998 #text_change 16-Jul-1999
 C:Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116; C4
 R:Flanagan, J.G.; Rabbitt, T.H.
 EMO J. 1, 655-660, 1982
 A:Title: The sequence of a human immunoglobulin epsilon chain constant region gen
 A:Reference number: A22771; MUID:84236029; PMID:6234164
 A:Accession: A22771
 A:Molecule type: DNA
 A:Residues: 1-428 <FLA>
 A:Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:9185035
 R:Ueda, S.; Nakai, S.; Nishida, Y.; Hiseajima, H.; Honjo, T.
 EMO J. 1, 1539-1544, 1982
 A:Title: Long terminal repeat-like element flank a human immunoglobulin epsilon pseudo
 A:Reference number: A23195; MUID:84207910; PMID:6327276
 A:Accession: A23195
 A:Molecule type: DNA
 A:Residues: 2-428 <UED>
 A:Cross-references: GB:J00222; NID:9184755
 R:Zhang, K.; Saxon, A.; Max, E.E.
 J. Exp. Med. 176, 233-243, 1992
 A:Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing
 A:Reference number: PH1214; MUID:92308839; PMID:1613458
 A:Accession: PH1214
 A:Molecule type: DNA
 A:Residues: 320-428 <ZHA>
 A:Cross-references: EMBL:X63693; GB:S18668; NID:932987
 R:Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasada, R.; Igatahi, K.; Kikuchi, M.; Sug
 Nucleic Acid Res. 11, 719-726, 1983
 A:Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon ch
 A:Reference number: A93491; MUID:83168897; PMID:6300763
 A:Accession: A93491
 A:Molecule type: mRNA
 A:Residues: 1-428 <SEN>
 A:Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:9185035
 R:Max, E.E.; Batey, J.; Ney, R.; Kirsch, I.R.; Leder, P.
 Cell 29, 691-699, 1982
 A:Title: Duplication and deletion in the human immunoglobulin epsilon genes.
 A:Reference number: A90824; MUID:83001945; PMID:6288268
 A:Accession: A90824
 A:Molecule type: DNA
 A:Residues: 1-358, 'L', 360-428 <MAX>
 A:Cross-references: GB:J00222; NID:9184755
 A:Note: this sequence difference may be due to polymorphism
 R:Benich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.
 in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-
 A:Reference number: A94418
 A:Accession: A94418
 A:Molecule type: protein
 A:Residues: 'GAWTL', 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 1
 A:Experimental source: myeloma protein Nd
 R:Kenten, J.H.; Molgaard, H.V.; Houghton, M.; Derzhavire, R.B.; Viney, J.; Bell, L.O.; C
 Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
 A:Title: Cloning and sequence determination of the gene for the human immunoglobulin ep
 A:Reference number: A93933; MUID:83065234; PMID:6815656
 A:Accession: B93933
 A:Molecule type: mRNA
 A:Residues: 1-40, 68-114, 427-428 <KEN>
 A:Cross-references: GB:L00022; NID:9185035
 R:Ikeyama, S.
 FEBS Lett. 224, 306-310, 1987
 A:Title: Purification and characterization of a recombinant human IGE Fc-epsilon fragmen
 A:Reference number: S02438; MUID:88083554; PMID:3121387
 A:Accession: S02438
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 98-352 <IKE>
 R:Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.

```

J. Biol. Chem. 269, 4566-462, 1994
A>Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces A:Reference number: J83116; MUID:94103254; PMID:8276835
A:Accession: J83116
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 320-428 <ZH2>
A:Experimental source: myeloma U266-derived cell line AF-10 R.Hellman, L. Eur. J. Immunol. 23, 159-167, 1993
A>Title: Characterization of four novel epsilon chain mRNA and a comparative analysis o A>Note: sequence extracted from NCBI backbone (NCBIN:141701, NCBIP:141702)
R.Hellman, L. Eur. J. Immunol. 23, 159-167, 1993
A:Reference number: A46536; MUID:93122085; PMID:8419166
A:Accession: C46536
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-426 <HEL>
A:Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:125297)
A:Note: sequence extracted from NCBI backbone (NCBIP:125297)
A:Accession: D46536
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-391 <HE2>
A:Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:125299)
A:Accession: A46536
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 401-428 <HE3>
A:Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:123483)
C:Genetics:
A:Gene: GDB:IGHF
A:Cross-references: GDB:119335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterotrimeric subunit consists of two identical light (Lap hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotrimer; immunoglob F:22-87/DNAin: immunoglobulin homology <IM1>
F:128-195/DNAin: immunoglobulin homology <IM2>
F:232-301/DNAin: immunoglobulin homology <IM3>
F:338-407/DNAin: immunoglobulin homology <IM4>
F:14/Disulfide bonds: interchain (to light chain) #status predicted
F:15-105,29-85,135-199,229-299,345-405/Disulfide bonds: #status predicted
F:121,49,99,146,252,275/Binding site: carbohydrate (Asn) #status experimental
F:121,209/Disulfide bonds: interchain (to heavy chain) #status predicted

Query Match          10.4%; Score 356; DB 1; Length 428;
Best Local Similarity   26.6%; Pred. No. 1e-13; Indels 102; Gaps 18;
Matches    118; Conservative      69; Mismatches     154;

QY       267 TODEKL-----OMGKKLPILHT-----LFOLPOVAGSGL--TLAEAKTGTI 308
        |||:||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       3 TGSSVFPLTRCCKNIPSATSVTLGCIALGYFFBPVVWTDTSINGTTMTLATITLT 62
QY       309 ---HGEVLVMKRAIQLOKNIICEV-WGPISPKMLSLIKENKAAYSKRE---KPVAWL 361
        ::::||||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       63 SGHVATISLLTVSGAMAKMFCTCRVAHPSTDW----VDNKFSVCSDRPFPPTVKIL 117
QY       362 N-----DEAGMGOCILL-----DSGOVTLESNIKULPWTGSTPPPCAPE 400
        ::||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       118 QSCDCGGHFPTTIQLDLAVSYTPGITINTIWLDGDGVMDV-----ISTASTOEGELAS 173
QY       401 PKS-----CDKTHTTC-----PELLGSPVFLFPPKPDTL 430
        :||:|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       174 TQSSELTLSOKKMLEDRTTYTCQTYYQGHTFEEDSTTKCADSNR---GISAVYSRSRPSFD-L 229
QY       431 MISRTPEVTCVVADVSHEDPEVKRMNVYDGVFNAAKTCKPREEOYNSTRVVSYLVIHQ 490

```

```
Db      FIRKSPITTLCLVLDLAPSKGVNLTWSPAAGKPVDHSTRKEKHQNGILTYTSTLPVGIR   289
Oy      DMLNGKEYCKRKNALPAPIEKTISKAKGQPREFOVYTL-----PPSRDELTKNQSVLT    545
Db      DWIEGETYCQRHTPHPLPRALMRSTTTKSGRPAAAEVAFAFATPEMPGRSDKRT-----LA  344
Oy      CLVKGFYSDDIAVEAESN--GGPENNYKITTPVLDSOSFLYLKYSLTYDKSNQGNNFS     603
Db      CLIQFMEDISVOQLNHNEVOLPDARHSSTOPRKTKSGSFVFPSRLRYRAEMEQDEPI    404
Oy      CSVMHEALTHMYT-OKSLSLSPG 625
Db      CRAVEHAASPGCTQVRRAVSANPG 427

RESULT 63
MERM
Ig mu chain C region, membrane-bound form - rabbit
C|Species: Oryctolagus cuniculus (domestic rabbit)
C|Date: 04-Dec-1986 #sequence_revision 30-Jun-1991 #text_change 23-Aug-1997
A|Accession: A02165; A02164
J.|Bernstein, K.E.; Alexander, C.B.; Reddy, E.P.; Mage, R.G.
A.|Title: Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of V-Haz
A.|Reference number: A02164; MUID:84088930; PMID:6418803
A.|Contents: a2 allotype
A|Molecule type: mRNA
A|.Residues: 433-479 <BE2>
A|Accession: A02164
A|Molecule type: mRNA
A|.Residues: 1-438,'GKPTLVNSLIIMSDTAETCY' <BER>
A|.Note: The sequence of residues 1-438 was assumed to be identical with the correspondi
C|Complex: An immunoglobulin heterotetramer subunit consists of two identical light ligh
Chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1
C|Superfamily: Immunoglobulin C region; immunoglobulin homology
C|.Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglob
F.121-92/DNA: immunoglobulin homology <IMM1>
F.130-202/DNA: immunoglobulin homology <IMM2>
F.242-310/DNA: immunoglobulin homology <IMM3>
F.348-420/DNA: immunoglobulin homology <IMM4>
F.433-479/DNA: carboxyl-terminal <CTS>
F.114/DNA:Disulfide bonds: interchain (to light chain) #status predicted
F.127-90,137-200,249-308,356-418/DNA:Disulfide bonds: #stratus predicted
F.46,114,212,261,277,284/Binding site: carbohydrate (Aan) (covalent) #statue predicted
F.129/Disulfide bonds: interchain (to heavy chain) #status predicted
F.129/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match          10.4%; Score 355.5; DB 1; Length 479;
Beat Local Similarity 24.8%; Pred. No. 1.3e+13;
Matches 122; Conservative 81; Mismatches 176; Indels 113; Gaps 21;

Oy      FPL-----AFIVFKLTSSGEIWMQAERASSKSWTFPLKXKEVS VKAYTDQPKLGKK       277
        :||               ||           |:||             ||         ||
Db      YPLVSCBGLTDGNLVAMGCARDLFLLSSVTFSW-SFK-NNSSEISSRVTRFFPVAVRCIDK   66
Oy      -----LPHHLTPLPOL-----POYA 292
Db      YMATSQVLVPESKDYLQGTTEHYLVCKVQHSHNSNRDLRSFPVDSLPPNVASFIPRDSFS   126
Oy      GSIGNLTIALLEAK-TGKLHQEVNLVVMTAQ-----LOKNLTCEEWGPTSFKLM,L,KLKE   345
        ||||              :::::            ||||||
Db      GSGRRKRSLIGQAGFGSPKOJSVSMWRDQCVCVESGYLTKPYEAELTKKGAGRPAFTSISSLTG  186
Oy      NKAERVSKREKRPVANVLPDAQMOCLLSDSQVILLNESIKYLPWTGSTFVPCCRAPBKSCD   405
        |:                ::          ::||
Db      ITESD-----WL-----SOELYTCRDVDRG-IFFDXKVSVMSSECST-----TSPP---  225
Oy      KTHPCBELLGGSVFLFRPKRKDTLMTSRPREVLCVVWDVHDHPYKFNMYYUGVEVHN      465
        ||||              |||||         |||||
Db      GIQVFPPIAPSADT-FUSKSAKRLICLVLYDILTITYG-SLNISM-----ASHN      268
```

```

QY 466 AKT-----KPREQNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIETKISKAG 520
D 269 GKRLDTHMNTESHHPNATFSAMEBASVCADWESGQFTCTVHADLPPLKATIKSKSE 328
QY 521 QPRE-POVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWEMESNGQP--ENNYKTPPV 576
D 329 VAGHPRAVYVLPAPAREGLVRESATYTCVKGSPADVPFQMOGRQPLSSDKYVTSARA 388
QY 577 LD--SDGSFPLYSKLTVDKSRMOQGVFSCSVMEALHNHYTKSLSPGLQDPTCAE 634
D 389 PEQADGLYFTHSTLTVTEEDMNSGETFTCVGHEALPHWTERTVDKS-----TEGEV 443
QY 635 AODELDGLMTT 646
D 444 AEEEGFENIMWT 455

RESULT 64
Ig epsilon-chain - chimpanzee (fragment)
C/Species: Pan troglodytes (chimpanzee)
C/Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C/Accession: I36948
R:Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A/Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orangutan
A/Reference number: I36948; MUID:87147196; PMID:3103123
A/Accession: I36948
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-426 <RES>
A/Cross-references: GB:M15398; NID:g176797; PIND:AAA3416.1; PID:g176798
C/Genetic:
A/Intons: 103/1; 209/1; 317/1
C/Superfamily: immunoglobulin C region; immunoglobulin homology
F/336-405/Domain: immunoglobulin homology <IMM>

Query Match 10.4%; Score 353.5; DB 2; Length 426;
Best Local Similarity 26.4%; Pred. No. 1.5e-13;
Matches 110; Conservative 68; Mismatches 162; Indels 77; Gaps 15;

QY 228 TVEKLTGSGELMW-----QAERASSKSMI--FPDLKKEVSVKRVTPDQLQWGGK 277
D 67 TISLITVSGA--NAKQMFTRVATPSTBVDNKTSVCSRDPTPVKVLQSSCGGGH 124
QY 278 LPLHLTLPQALPOYA--GSGNLTLALEAKTGKHOEVNLVVMRAITOLQKXLTCEWGGTSP 336
D 125 FPPITQLCLVSGVTPGTINITWLEDOV--MDVDLSTASATQEGE-----LASTGS 174
QY 337 KLMISLKLKNEKAKVSKREKRPVWVNLNPAEAGMOCILSDSQVLLSENIKVLPTWSTPVPC 386
D 175 ELTLISQK-----HMLSD--RTYTCQVTVYGGTFEDSTKK-----C 207
QY 397 PAPEPKSCDKHTHPCPELLGSPVFLPPKPRDTLMISRTPEVTCVAVDVSHEDPEVKFNW 456
D 208 ADSNPR-----GVSAVLSPRPSPD-LFIKSPITICLVVDLAPSKGTAVLTW 253
QY 457 YVDGVEVNAKTKPREBOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIETKIS 516
D 254 SRASGKGVNHNSTRKQKORNGITLVSTLPGVGRDWIEGEVQGRVTHPHLPALVYST 313
QY 517 KAKQPREPOVYTL-----PPSRDELTKNOVSLTCLVKGFPSPDIWEMESN--GQPENN 569
D 314 KTSQPRAPAEVYATPREGPSRDKRT-----LACLQNTMPEDISVQMLHNEVOLPDA 368
QY 570 YKTPPVLDSGSEFPLYSKLTVDKSRMOQGVFSCSVMEALHNHYT--OKSLSPG 625
D 369 HSTTQPKTKSGCFVFSRLVETPAEWQKDEFCRAVNHAPSQSVQTVTSVNP 425

RESULT 65
MHMUBT
Ig mu heavy chain disease protein (Boc) - human

```

```

C/Species: Homo sapiens (man)
C/Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jul-1999
C/Accession: A02163
R:Barnikol-Watanabe, S.; Mihaesco, E.; Mihaesco, C.; Barnikol, H.U.; Hilschmann, N.
Hope-Seyler's Z. Physiol. Chem. 365, 105-118, 1984
A/Title: The primary structure of mu-chain-disease protein BOT. Peculiar amino-acid seq
A/Reference number: A02163; MUID:84184186; PMID:4425189
A/Accession: A02163
A/Molecule type: protein
A/Residues: 1-391 <BAR>
C/Comment: This protein has no V region homology or CH1 region.
C/Genetic:
A/Genes: GDB:IGHM
A/Cross-references: GDB:120086; OMTM:147020
A/Map position: 14q32.33-14q32.33
A/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin; transmembrane pr
F/1-41/Domain: pre-C <VAR>
F/43-391/Domain: Ig mu chain C region, secreted form <IGM>
F/65-137/Domain: immunoglobulin homology <IMM1>
F/175-243/Domain: immunoglobulin homology <IMM2>
F/282-353/Domain: immunoglobulin homology <IMM3>
F/147,210,217,378/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.3%; Score 352.5; DB 1; Length 391;
Best Local Similarity 27.8%; Pred. No. 1.5e-13;
Matches 84; Conservative 61; Mismatches 110; Indels 47; Gaps 11;

QY 332 GPTSPMLSLKLENKAKVSKREKRPVWVNLNPAEAGMOCILSDSQVLLSENIKVL--PT 389
D 110 GPTTKVNTSTLTIKSD-----WL---GQSMFTCRVDHRG-LTRQVASSMGCPD 155
QY 390 WSTPVPAPAPKSCDKHTHPCPELLGSPVFLPPKPRDTLMISRTPEVTCVAVDVSHD 449
D 156 QDTA-----RVFAIPPS--FASIFLTSTKTLCLVTDLTVD 191
QY 450 PEYKFNWYVDGVEVNAKTKPREBOYNSTYRVSVLTVLHODMLNGEKYCKVSNKALPA 509
D 192 -STTISWTRQDEAVKTHTNISESHNATFSVGEASICEEDMDSERTCTVTHIDLP 250
QY 510 PIEKTSKAKGP-REPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWEMESNGQP- 566
D 251 PLKQTSIRKGVALLRPDLYLLPPAREQLNLRESATITCLVGFSDVDFVQMOGRQPL 310
QY 567 -ENNYKTPPVLD--SDGSFPLYSKLTVDKSRMOQGVFSCSVMEALHNHYTKSLSP 623
D 311 SPEKYTSAPMEPQAPGRYFAHSILTVSEEMNTGETYTCVVAHEALPNRYTERVDS 370
QY 624 PG 625
D 371 TG 372

RESULT 66
MHMUM
Ig mu chain C region, membrane-bound splice form - human
C/Species: Homo sapiens (man)
C/Date: 29-Jan-1993 #sequence_revision 23-Aug-1997 #text_change 22-Jun-1999
C/Accession: S16510; S09357; S16556; B26243; A02167
R:Dorai, H.; Gillies, S.D.
Nucleic Acids Res. 17, 6412, 1989
A/Title: The complete nucleotide sequence of a human immunoglobulin genomic C-mu gene.
A/Reference number: S09357; MUID:89366690; PMID:2505237
A/Accession: S16510
A/Molecule type: DNA
A/Residues: 433-473 <DOR1>
A/Cross-references: EMBL:X14939
A/Note: the authors translated the codon AAC for residue 445 as Met
A/Note: the sequence of residues 1-432 was assumed to be identical with the correspondi
A/Molecule type: DNA
A/Residues: 1-432; GKPLTVNVLVMSDPTAGTCY' <DOR2>
A/Cross-references: EMBL:X14940

```


A>Note: the authors translated the codon AAT for residue 16 as Met
A>Note: secreted splice form
R:Doral, H.
submitted to the EMBL Data Library, April 1989
A:Reference number: S16656
A:Accession: S16656
A:Molecule type: DNA
A:Residues: 1-39, 'L', '41-432, 'GKPTLVNLSLWSDTACTY' <DOR3>
A:Cross-references: EMBL:X14940
R:Rabbits, T.H.; Forester, A.; Mlstein, C.P.
Nucleic Acids Res. 9, 4509-4524, 1981
A:Title: Human immunoglobulin heavy chain genes: evolutionary comparisons of C-mu, C-de1
A:Reference number: A26243; MUID:82059479; PMID:6795593
A:Accession: B26243
A:Molecule type: DNA
A:Residues: 433-436, 'N', '438, 'E', '440-447, 'T', '449-473 <RAB>
A:Cross-references: GB:K01310; NID:G184715; PIDD:AA559422.1; PID:G184734
C:Comment: During differentiation, B lymphocytes switch from expression of membrane-bound
B.
C:Genetics:
A:Gene: GDB:IGHM
A:Cross-references: GDB:120086; OMTM:147020
A:Map position: 14q32.33-14q32.33
A:introns: 1/1; 105/1; 217/1; 323/1; 433/1; 471/3
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. The IGH subunits associate into disulfide linked pentamers.
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; glycoprotein; heterotrimer; immunoglobulin; membrane
F:127-199/DNA: immunoglobulin homology <IMM1>
F:127-199/DNA: immunoglobulin homology <IMM2>
F:127-305/DNA: immunoglobulin homology <IMM3>
F:127-305/DNA: immunoglobulin homology <IMM4>
F:433-473/DNA: carboxyl-terminal #status predicted <CTS>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:28-88,134-197,244-303,351-413/Disulfide bonds: #status experimental
F:46,209,272,279/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:214/Disulfide bonds: interchain (to heavy chain) #status experimental
F:291/Disulfide bonds: interchain (to mu chain in another subunit) #status experimental

Query Match 10.3%; Score 351.5; DB 1; Length 473;
Beat Local Similarity 27.0%; Pred. No. 2.2e-13;
Matches 87; Conservative 64; Mismatches 120; Indels 51; Gaps 12;

332 GPTSPKMLSLKLEKAKVSKREKPVVNLPEAGMOCCLSDSGQVLE-SNKKVPTW 390
172 GPTTYKVTSTLTKESSD-----WL---GSGMFCRDVHRLGTFQGNASSMCVPDQ 218
391 STVPCCAPBPBKSCDKHTCEPLLGGPSVFLPPKPKDMLISRTPEVTCVVVSHEDP 450
219 DTAL-----RVFAIPPS-FASIFLTSTKLTCLVTDLTYYD- 253
451 EVKRWVYDGVENNAKTPREBOGNSYRVVSLTVLHQMNGKEYCKVSKKALPAP 510
254 SVTISWRQNGEAVKTHNISESHPNFSAVGEASIEDDMNGSERPTCTVHTDLPSP 313
511 IEKTSKAKGP-REPOVYTLPPSRDELTKNOVSLTLVNGFPPSDAVWESNGQP-- 566
314 LKQITSRKGVALHRPDVYLLPPAREQLNRESATITCLVTFSPADVFQVMORGPLS 373
567 ENNKVTPPVLD--SDGSFFLYSKLTVKSRMOQGNVSSCVMEALHNHTQKSLSP 624
374 PEKVYTSAPMEPPQAPGKGFYFAHSILITVSEEMNGERTYTC-VAHEALPNRTKTVDSK- 431
625 GLQDDEFCAEQDGLDGLMTT 646
432 -----TEGEVSADEEGFENLMAT 449

RESULT 67
MHMS
1g mu chain C region, secreted form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Oct-1980 #sequence_revision 31-Oct-1980 #text_change 16-Aug-1996

C:Accession: A02166; A26239; A26240; B02039
R:Kawakami, T.; Takahashi, N.; Honjo, T.
Nucleic Acids Res. 8, 3933-3945, 1980
A:Title: Complete nucleotide sequence of mouse immunoglobulin mu gene and comparison w
A:Reference number: A02166; MUID:81076590; PMID:6255422
A:Accession: A02166
A:Molecule type: DNA
A:Residues: 1-455 <KAW>
A:Cross-references: GB:J00443
A>Note: the sequence was determined from the germline gene
R:Goldberg, G.I.; Vanu, E.F.; Zrolka, A.M.; Blattner, F.R.
Gene 15, 33-42, 1981
A:Title: Sequence of the gene for the constant region of the mu chain of Balb/c mouse
A:Reference number: A26239; MUID:82051295; PMID:6795090
A:Accession: A26239
A:Molecule type: DNA
A:Residues: 1-455 <GOL>
A>Note: the sequence was determined from the germline gene
R:Auffray, C.; Rougeon, F.
Gene 12, 77-86, 1980
A:Title: Nucleotide sequence of a cloned cDNA corresponding to secreted mu chain of mou
A:Reference number: A26240; MUID:81165562; PMID:6260591
A:Accession: A26240
A:Molecule type: RNA
A:Residues: 1-225, 'N', '227-257, 'S', '259-367, 'K', '369-455 <AUF>
R:Kehry, M.; Sibley, C.; Fuhrman, J.; Schilling, J.; Hood, L.E.
Proc. Natl. Acad. Sci. U.S.A. 76, 2932-2936, 1979
A:Reference number: A26241; MUID:79223904; PMID:111247
A:Contents: annotation; MOPC 104E
A>Note: This sequence has been revised in reference A02039. Carbohydrate binding sites
R:Kehry, M.R.; Fuhrman, J.S.; Schilling, J.W.; Rogers, J.; Sibley, C.H.; Hood, L.E.
Biochemistry 21, 5415-5424, 1982
A:Title: Complete amino acid sequence of a mouse mu chain: homology among heavy chain c
A:Reference number: A02039; MUID:83075344; PMID:6816276
A:Contents: MOPC 104E
A:Accession: B02039
A:Molecule type: protein
A:Residues: 1-77, 'N', '79-100, 'Q', '102-225, 'N', '227-257, 'T', '259-367, 'K', '369-455 <KEH>
C:Genetics:
A:introns: 1/1; 106/1; 219/1; 325/1
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgM and IgG, the subunits associate into 1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotrimer; immunoglob
F:129-201/DNA: immunoglobulin homology <IMM1>
F:129-201/DNA: immunoglobulin homology <IMM2>
F:129-307/DNA: immunoglobulin homology <IMM3>
F:129-307/DNA: immunoglobulin homology <IMM4>
F:436-455/DNA: carboxyl-terminal <CTS>
F:14/Disulfide bonds: interchain (to light chain) #status predicted
F:28-89/Disulfide bonds: #status experimental
F:46,211,243,281,442/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:136-199,246-303,351-415/Disulfide bonds: #status predicted
F:216,454/Disulfide bonds: interchain (to heavy chain) #status predicted
F:291/Disulfide bonds: interchain (to mu chain in another subunit) #status predicted

Query Match 10.2%; Score 348; DB 1; Length 455;
Beat Local Similarity 22.7%; Pred. No. 3.3e-13;
Matches 122; Conservative 77; Mismatches 179; Indels 160; Gaps 18;

144 ESPPGSSPVQGNRP-RGKTIQGGKTLISVQLELDSDGTYTCVTLQNKVYEFKIDIVL 202
3 QSPFNVPPLVSCSPSLSDKXILVMGCLARDFLSTISFTN----- 43
203 AFGKASIVYKKGGEVFPFPLAFYTKULTGSGELMWQARASSSKSWITFDLKNKEVS 262
44 -----YQNTVEVIQ-----GIRTFPLTGGKYLA 68
263 VKRVTQDPKLOM-----GKKLPLHLTLPOALPOYAGSGNLTALAEAKTKGL 308
69 TSGVLSPPSKILSGSDYLVCKIHYGKKNRDLVVP-----AVAEW 110

A:Cross-references: EMBL:V00563; NID:G33454; PIDN:CA23826.1; PID:9825684
 R:Mihaesco, E.; Barnikol-Matanabe, S.; Barnikol, H.U.; Mihaesco, C.; Hilschmann, N.
 Eur. J. Biochem. 111, 275-286, 1980
 A:Title: The primary structure of the constant part of mu-chain-disease protein BOT.
 A:Reference number: A02162; MUID:81066716; PMID:6771162
 A:Contents: Mu-chain-disease protein Bot and revisions to sequence of Gal
 A:Accession: A02162
 A:Molecule type: protein
 A:Residues: 117, 'BPS', '22', 'T', '24-82', 'N', '84-90', 'Z', '92-93', 'B', '95', 'B', '97-144', 'E', '146-162', 'E'
 A:Accession: B02162
 A:Molecule type: protein
 A:Residues: 100-144, 'E', '146-162', 'E', '164', 'E', '166-214', 'G', '216-262', 'D', '264-295', 'D', '297-414',
 A:Note: all four combinations of the 191-ser/Gly and 215-val/Gly polymorphisms have been
 R:Matanabe, S.; Barnikol, H.U.; Horn, J.; Bertram, J.; Hilschmann, N.
 Hoppe-Seyler's Z. Physiol. Chem. 354, 1505-1509, 1973
 A:Title: The primary structure of a monoclonal Igm-immunoglobulin (macroglobulin Gal), I
 A:Reference number: A02064; MUID:75059123; PMID:4803843
 A:Contents: annotation; Waldenstrom's macroglobulin Gal
 A:Note: this sequence has been revised in reference A02162
 R:Putnam, F.W.; Florent, G.; Paul, C.; Shinoda, T.; Shimizu, A.
 Science 182, 287-291, 1973
 A:Title: Complete amino acid sequence of the mu heavy chain of a human Igm immunoglobulin
 A:Reference number: A02088; MUID:74005511; PMID:4742735
 A:Contents: annotation; Waldenstrom's macroglobulin Ou, sequence, disulfide bonds, and c
 A:Note: this sequence differs from that shown at a number of positions; this sequence ha
 C:Comment: During differentiation, B lymphocytes switch from expression of membrane-bound
 B
 C:Genetics:
 A:Gene: GDB:IGHM
 A:Cross-references: GDB:120086; OMIM:147020
 A:Map position: 14q32.33-14q32.33
 A:Introns: 1/1, 105/1, 217/1; 333/1; 433/1
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
 hain disulfide bonds. The Igm subunits associate into disulfide linked pentamers.
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: alternative splicing; glycoprotein; heterotetramer; immunoglobulin
 F:21-90/Domain: immunoglobulin homology <IMM1>
 F:127-199/Domain: immunoglobulin homology <IMM2>
 F:237-305/Domain: immunoglobulin homology <IMM3>
 F:343-415/Domain: immunoglobulin homology <IMM4>
 F:432-452/Domain: carboxyl-terminal <CTS>
 F:14/Disulfide bonds: interchain (co light chain) #status experimental
 F:28-88,134-197,244-303,351-413/Disulfide bonds: #status experimental
 F:46,209,212,219,439/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:214,451/Disulfide bonds: interchain (co heavy chain) #status experimental
 F:291/Disulfide bonds: interchain (co mu chain in another subunit) #status experimental

Query Match 10.2% Score 347; DB 1; Length 452;
 Best Local Similarity 27.6% Pred. No. 3.8e-13;
 Matches 83; Conservative 62; Mismatches 110; Indels 46; Gaps 11;

Qy 332 GPTSPKMLSLKLNKKAQVSKREKPVVNLNPEAGMQLSDSGQVLE-SNIKVLPTW 390
 Db 172 GPTTYKVTSTLTIESD-----WL-----GSMFRCRDHGLTFQQAASMCVPDQ 218
 Qy 391 STVPVCPAPKPSCKDTHTCPELLGGPSVFLPFPKPKDTLMIISRTPEVTCVAVDSHPD 450
 Db 219 DTAL-----RVFAIPPS-FASIFLTKSTKLTCLVDTLTYD- 253
 Qy 451 EVKKNWYVDGVEVNAKTKPREEQNSTYRVSVLTVLHODMLNGKEKCKVSKALPAP 510
 Db 254 SVTISWRQNGEAVKTHNISHPNATFSAVGSASICEDMDNNGERTCTVTHTDLPSP 313
 Qy 511 IEKTSKAKGP-REPQVYTLPSRDELTKNOVSLTCLVGFYPSDIAVEMSGOP-- 566
 Db 314 LKQTSIRKGVALLRPDVTLLPAPAEQNLRESATITCLVTGFSPADVFMQMGQPLS 373
 Qy 567 ENNYKTPPVLD--SDGSFPLYSKLTVDKSRWQGNVFSQVMEALNHYTKSLSLSP 624
 Db 374 PEKYVTSAPMDEPQAPGRYFAHSILITVSEEMNGETVTC-VAHEALPNRTERTVSKST 432
 Qy 625 G 625

Db 433 G 433

RESULT 70
 EHMS
 Ig epsilon chain C region (version 2) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996
 R:Shida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.
 EMBO J. 1, 1117-1123, 1982
 A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison w
 A:Reference number: A90966; MUID:84236092; PMID:6329728
 A:Accession: A02145
 A:Molecule type: DNA
 A:Residues: 1-423 <ISH>
 A:Note: the sequence was determined from the germline gene
 C:Genetics:
 A:Introns: 91/1; 199/1; 307/1
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:16-77/Domain: immunoglobulin homology <IMM1>
 F:115-183/Domain: immunoglobulin homology <IMM2>
 F:220-288/Domain: immunoglobulin homology <IMM3>
 F:325-386/Domain: immunoglobulin homology <IMM4>
 F:23-75,122-181,227-286,332-394/Disulfide bonds: #status predicted
 F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1% Score 346.5; DB 1; Length 423;
 Best Local Similarity 26.5% Pred. No. 3.7e-13;
 Matches 118; Conservative 74; Mismatches 154; Indels 99; Gaps 21;

Qy 223 PPLAFTEKLTGSGELMWQBRASSSWITPDLKNE--VSVKRYVQDPTLGWKKLPL 280
 Db 29 FPNPVYIT-----WYSD--SLNISTVNFPAISGELKVTTSQVTSWGK--SAKPYTC 75
 Qy 281 HLTLPOALPOYAGSGNLTALBAKTKLHOEVLVNR-----ATOLQKNLTCEVWGPSTP 336
 Db 76 HVTHPSPFNE-----SRTILVRYPTHTSLSPWMSYHRCBPMAFHTIQYCFYGHILN 130
 Qy 337 KLMISLLENKE-----AKSKREKPVVNLNPEAGMO-----CLSDSQVL 379
 Db 131 DVSVSMIDREITDITLAQTVLKEEGKLASTGSKLITBQWMSSTFCRYVSGVDY 190
 Qy 380 LESNIXVLPTWSTVPVCPAPKPSCKDTHTCPELLGGPSVFLPFPKPKDTLMIISRTPEVT 439
 Db 191 LAHTR-----CPDHEPR-----CAITYLIPPSFLD-LYQNGAKLT 226
 Qy 440 CVVVDV-SHPDPEVKFN-----WYVDGVEVNAKTKPREEQNSTYRVSVLTV 487
 Db 227 CLVVDLESEKNVNTNNOEKKTSVASAQW---TKHNN-----NATTSITSLPV 273
 Qy 488 LHODMLNGEKYCKKCVSKALPAPIEKTISKAK-GQPEPQVYTLPSRDELTKNOVSLTC 546
 Db 274 VAKWDIEGYGQCVVDPPDPKPKIVRSITLPQVQSRAPVYVFPPEER-SEKRTLTTC 332
 Qy 547 LVKGFPSPDIAYVESNGQPENN--YKTPPVLDSDS---FFLYSKLTVDKSRWQGNV 601
 Db 333 LIQNFPEPDISVQWLGSKLISNSQHSITTP-LKSNNGSGFFLPSHLEVAKTLMTORKQ 391
 Qy 602 FSCSVMEALHN-HYTKSLSLSPG 625
 Db 392 FTQGVHEALQKPKLEKTIISTSLG 416

RESULT 71
 MAHY
 Ig mu chain C region - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996

Best Local Similarity 24.8%; Pred. No. 4.4e-13;
Matches 117; Conservative 78; Mismatches 168; Indels 108; Gaps 20;

```

Qy 223 FPL-----AFTVEKLTGSGELMWOERASSKSWITPDLKNEVSVKRVTDPKLQNGKK 277
    |||
Db 9 YPLSCGALTDGNLVAMGCLARDPLSSVTFSSW-SFK-NNSEISSRTVRTPVVKGDGK 66
    |||
Qy 278 -----LPLHLTPQAL-----POYA 292
    |||
Db 67 YMATSOVLVPSKDVLOQTEEYLVCKVQHSNNDLRVSFPVDSLEPPNVSVFIPRDSFS 126
    |||
Qy 293 GSGNLTLLLEAK-TGKLHOEVNLVVMRATQ-----LQKNLTCEVWGPTSPKMLSLKLE 345
    |||
Db 127 GSGRKSLICQAGTGFSPKQISVSWLRDQGVESGVLTKEPAEATKAGAPRTFSSMLT 186
    |||
Qy 346 NKEAKVSKREKRPVWVLANPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCPADPEKSCD 405
    |||
Db 187 ITESD-----WL-----SQSLYTCRVDRHG-IFEDKNVSSSECSST-----TPSP----- 225
    |||
Qy 406 KTHCPPELLGSPVFLPPPKKDTLMTSRTEVTCVAVDVSHEDPEVKENYVDGVEVHN 465
    |||
Db 226 -----GIQVFPPIAPSPADT-FLSKSARLCLVTDLTYYG-SLNTISW-----ASHN 268
    |||
Qy 466 AKT-----KPREEQNSTYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEXTISKAKG 520
    |||
Db 269 GKALDTHNITESHFNATFSAMGASVCAEDMESGEOTCTVTHTADLPFLKHTISKSRE 328
    |||
Qy 521 QPRE-POVYTLPPSRDELTKNOVSLTCLVKGFPSPDIWESNGQPN--ENNYKTPPV 576
    |||
Db 329 VAKHPAYVLPAREQLVRESATVTCVKGFSPADVFVQMOGRGLSDKVTYSABA 388
    |||
Qy 577 LD--SDGSFLYSLKTLVDKSRMOQGNVFCGVMEALHNHTQKSLSLSPG 625
    |||
Db 389 PEQAPGLYFTHSTLTYTEEDMNSGETFTCVGHEALPHMTERTVDKSTG 439
    |||

```

RESULT 74

IG heavy chain - nurse shark
C/Species: Ginglynostoma citratum (nurse shark)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C/Accession: 150731
R/Vazquez, M.; Mituki, N.; Flajnik, M.F.; McKinney, E.C.; Kasaahara, M.
Mol. Immunol. 29, 1157-1158, 1992
A/Title: Nucleotide sequence of a nurse shark immunoglobulin heavy chain cDNA clone.
A/Reference number: 150731; MUID:92357056; PMID:1495502
A/Accession: 150731
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-577 <VAZ>
A/Cross-References: GB:M92851; NID:G213264; PIDN:AAA50817.1; PID:G213265
C/Genetic8:
A/Genes: IGH
C/Superfamily: immunoglobulin C region; immunoglobulin homology
F:469-539/Domain: immunoglobulin homology <IMM>

Query Match 10.1%; Score 345.5; DB 2; Length 577;
Best Local Similarity 23.9%; Pred. No. 6.3e-13;
Matches 155; Conservative 110; Mismatches 256; Indels 127; Gaps 29;

```

Qy 10 LLLVLAQLALPPAAGQGNKV-----LGKKDTEVELTCAQSQKSIQPMWKSNG--IKIL 62
    |||
Db 6 IFSLALLALPCVQSEITLQPEARTGPGSLSTC-----KTRGFVLGSSSMYWRV 60
    |||
Qy 63 GNCG-----SFLTKGPKLNDRADSRSLWDQGN--PPLIKMLKIEDSTYICEV 111
    |||
Db 61 PGQGLEWIVVYSSSMNNYPAIDRFPAK--DTSNIFALMBRSKIDDTAIYIC-- 115
    |||
Qy 112 EDQKEVQLVFLGTLANSDTLHLOGQSITLTLESPPGSSPVSQCRSPKGNIOGKTLV 171
    |||
Db 116 -----TRMSGYEYLGHSYGWQGTWVTATTATP--SSPLY-----GLVSSC 157
    |||
Qy 172 SQLELDQSGTWTCTVNLQNGKVEFKIDIVLAFQKASSIVYKKEGQVEFSFLAFVYK 231
    |||

```

158 QQGNIDSSVYIGCLAM-----DYSPIV-----ASVTMKKGGLITGVQTPSEVRN 203

```

Qy 232 LTGSGELMWOERASSKSWITPDLKNEVSVKRVTDPKLQNGKK 289
    |||
Db 204 KKGTYTL-----SSQALIESDACECDQISCEVRHSGSDKSTGM-----PCPDGFPALL 252
    |||
Qy 290 QVAGSGNLT-----LALFAKTKLHOEVNLVVMRATQ-OKNLTCGVWGPSPKMLSLKLE 345
    |||
Db 253 TVSSSEIEERKPAITVCSISDFRSK--SISVTMLKNGRSYDSGIFTSB-----VCEA 303
    |||
Qy 346 NKEAKVSKREKRPVWVLANPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPCPADPEKSCD 405
    |||
Db 304 NGNSVTSRLRVPAEMFDDAVYTC-----QVKYK--EVIQSWN-----ITGQVSECH 350
    |||
Qy 406 KTHCPPELLGSPVFLPPPKKDTLMTSRTEVTCVAVDVSHEDPEVKENYVDGVEVHN 465
    |||
Db 351 -----GYTAKILPP-FVEQVLLAEATVTLTCV--VSNLHSGVNFMTLOD----- 390
    |||
Qy 466 AKTKPREEQNS--TYRVSVLTVLHODMLNGEKYCKVSNKALPAPIEXTISKAK-GQ 521
    |||
Db 391 EKLKSIADSGEHSQALSKLDISTEAWLSEVEECVNHQYLPPLRDSIHKEHEN 450
    |||
Qy 522 PREPOVYTLPPSRDELTKNO-VSLTCLVKGFPSPDIWESNGQPN--NYKTPPVLD 578
    |||
Db 451 PLEBSVVLPTTEBELSAQRFSLTCLVGRFRREIVKKTITNDKPYNPENYKTEVTAL 510
    |||
Qy 579 SDG-SFPLYSKLTVDKSRMOQGNVFCGVMEALHNHTQKSLSLSPG 625
    |||
Db 511 SDNTSFPLYSLSLIAEEMASGASVSCVGHAIPLKTIINRTVDKSSG 558
    |||

```

RESULT 75

IG mu chain C region - rat (firearm)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 20-Feb-1995 #sequence_revision 30-Jan-1998 #text_change 21-Jan-2000
C/Accession: S25644
R/Parker, K.; Bugeon, L.; Soulliou, J.P.
submitted to the EMBL Data Library, September 1992
A/Reference number: S25644
A/Accession: S25644
A/Molecule type: mRNA
A/Residues: 1-343 <PAR>
A/Cross-References: EMBL:X68312; NID:956461; PIDN:CAA48392.1; PID:9818025
A/Experimental source: spleen
C/Genetic8:
A/Map position: 6
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin
F:234-305/Domain: immunoglobulin homology <IMM>

Query Match 10.0%; Score 341.5; DB 2; Length 343;
Best Local Similarity 29.5%; Pred. No. 5.6e-13;
Matches 96; Conservative 59; Mismatches 129; Indels 41; Gaps 12;

```

Qy 323 OKNLTCGVWGPSPKMLSLKENKAKVSKREKRPVWVLANPEAGMOCCLSDSGOVLLES 382
    |||
Db 19 KSLRICATNFSPKQITVSWLODKPKVSGFTTEPVTV-EAKGSRPQTYKIVSTLTITTES 77
    |||
Qy 383 ---NIKV-----LPTWSTPVPCPADPEKSCDKTHCPPELLGSPVFLPPKPKDTL 430
    |||
Db 78 DWLNLNFTCRVDHRLGTLTFKKNVSTCAASPT-----DILAF-----IPSPAD-I 124
    |||
Qy 431 MISTPEVTCVAVDVSHEDPEVKENYVDGVEVHNATKPREEQNSTYRVSVLTVLH 490
    |||
Db 125 FLTKSATLSGLVNTLAIYD-TLNISMSKSGEFLNTKTMESHNPOTFSAVGVASVCE 183
    |||
Qy 491 DWLNGEKYCKVSNKALPAPIEXTISKAKQPRE-----POVYTLPPSRDEL-TKNQVSL 544
    |||
Db 184 DMDNRKEFVCTVTHRDLPSPQKFKISK--PNEVAKHPAVVYLLPPARQQLIRRESATV 239
    |||
Qy 545 TCLVKGFPSPDIWESNGQPN--ENNYKTPPVLD--SDGSFLYSLKLTVDKSRMOQGN 600
    |||

```

Db 240 TCIVKFSRPAIDIVQWLGKQPLSSDKYVTSAMPBPGAGLYFTHSILTVIEEENNSE 299
 QY 601 VFSGSVMHEALNNHYTKSLSPG 625
 Db 300 TYTCVGVGHEALPMHVTERTVDKSTG 324

RESULT 76

MHDG
 Ig mu chain C region - dog (tentative sequence)
 C/Species: Canis lupus familiaris (dog)
 C/Date: 13-Jul-1981 #sequence revision 13-Jul-1981 #text_change 31-Mar-2000
 C/Accession: A93131; A94246; A02169
 R/McCumber, L.J.; Capra, J.D.
 M01. Immunol. 16, 565-570, 1979
 A/Title: The complete amino-acid sequence of a canine mu chain.
 A/Reference number: A93131; MUID:80077682; PMID:117299
 A/Contents: myeloma protein Moo
 A/Accession: A93131
 A/Molecule type: protein
 A/Residues: 1-177 <MCC>
 R/Wasserman, R.L.; Capra, J.D.
 Science 200, 1159-1161, 1978
 A/Title: Amino acid sequence of the Fc region of a canine immunoglobulin M: interspecies
 A/Reference number: A94246; MUID:78180587; PMID:653360
 A/Contents: Moo
 A/Accession: A94246
 A/Molecule type: protein
 A/Residues: 178-450 <MWS>
 C/Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 16 C/Superfamily: immunoglobulin C region; immunoglobulin homology
 C/Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F/20-89/Domain: immunoglobulin homology <IMM1>
 F/126-196/Domain: immunoglobulin homology <IMM2>
 F/234-302/Domain: immunoglobulin homology <IMM3>
 F/341-412/Domain: immunoglobulin homology <IMM4>
 F/430-450/Domain: carboxyl-terminal <CTS>
 F/14/Disulfide bonds: Interchain (to light chain) #status predicted
 F/27-87,133-194,348-410/Disulfide bonds: #status predicted
 F/50,206,269,276,437/Binding site: carbohydrate (asn) (covalent) #status predicted
 F/211,449/Disulfide bonds: Interchain (to heavy chain) #status predicted
 F/288/Disulfide bonds: Interchain (to mu chain in another subunit) #status predicted

Query Match 10.0%; Score 340; DB 1; Length 450;

Best Local Similarity 25.9%; Pred. No. 9.7e-13;

Matches 100; Conservative 76; Mismatches 142; Indels 68; Gaps 15;

QY 278 LPLHLTLPOAL-----POVAGSGN---LTLALEAKGCKLHGVNLLVMRATOLQKNT-- 327
 Db 102 LPWMLTLPPEVSGFTIPRDAFGSPKRSQILCOAGSPQVWSL--RDGKQIESGVTIN 159
 QY 328 -----CEVWGPTSPKMLSLKLENKEAKVSKREKPVVNLPEAGMOCCLSDSGVLLS 382
 Db 160 EVZAAZAKZSGPTTYKYVSMULTI-----QEDAL---SQSVFCXKHHNG-LTFQ 205
 QY 383 NIKVLPTWSTPVPCPAPEPKSCDKHTTCPELLGGSPVFLFPKPKDTLMSIRPEVTCV 442
 Db 206 NASMCTSDQPV-----GISFTIPPS-FASIFPTKSAKSLCV 243
 QY 443 VDVSHEDEPKVFNMYVGVGVHNAKTKPREQVNSTRVVSVLTVLHODMLNKEXYCKV 502
 Db 244 TDLATYV-STVISTREBNGALKTHINISHPNGTSAANGEATVCEEESGSEGFCTV 302
 QY 503 SNKALPAPIEKTSKAGQ-OPREPOVYTLPPSRDEL-TKNQVSLTCLVKGFYPSDIAVEN 560
 Db 303 THDLPVSLKQTSIRPQAVNHMPVYVLPSPRQGLRLRSATLSCLVTSYSPDVVQW 362
 QY 561 ESNQO--PENNYKTPPVLD--SDGSFLLSKLTVDKSRNQGVSCSMHEALNNHYT 616
 Db 363 VOKQQVPDPDSYVTSAMPBPGAGLYFAHSILTVSEENNAGETVTCVVAHESLPKRV 422

QY 617 OKSLSPG-----LQDDETCAR 634
 Db 423 ERSVDSKCKPVLVNVSLVLSPTAGZ 448

RESULT 77

B30503
 Ig gamma-2a chain C region (B5.7A12) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 31-Mar-1989 #sequence revision 30-Sep-1991 #text_change 16-Jul-1999
 C/Accession: B30503
 R/Gilmore, G.L.; Bard, J.A.; Birshnein, B.K.
 J. Immunol. 141, 1754-1761, 1988
 A/Title: DNA rearrangements affecting both variable and constant regions of Ig H chain.
 A/Reference number: A30503; MUID:88315788; PMID:2842402
 A/Accession: B30503
 A/Molecule type: mRNA
 A/Residues: 1-112 <GIL>
 A/Cross-references: GB:M21925
 A/Experimental source: myeloma cell line MPC11
 A/Note: the authors translated the codon GAG for residue 11 as Ser
 C/Genetics: 100/3
 A/Intons: 100/3
 C/Superfamily: immunoglobulin C region; immunoglobulin homology
 C/Keywords: immunoglobulin
 F/1-70/Domain: immunoglobulin homology <IMM>

Query Match 9.7%; Score 331.5; DB 2; Length 112;

Best Local Similarity 50.0%; Pred. No. 5.2e-13;

Matches 66; Conservative 17; Mismatches 22; Indels 27; Gaps 2;

QY 433 SRTPEVTCVNVVSHBDEPKFNWYVDGVGVHNAKTKPREQVNSTRVVSVLTVLHODM 492
 Db 1 SLTPKTCVNVVSEDDPVQLSWFNVNVEVTAQQTREEDYNSTIRRVSTLPIOHDM 60
 QY 493 LNKGEYKCKVSNKALPAPIEKTSKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFY 552
 Db 61 MGKKEKCKVNNKDLPAPIERTISKPGK-----SCSLTAMGLGW- 100.
 QY 553 PSDIAVENMSG 564
 Db 101 -----WTSNG 105

RESULT 78

S60266
 novel antigen receptor precursor - nurse shark
 C/Species: Ginglymostoma cirratum (nurse shark)
 C/Date: 10-Apr-1996 #sequence revision 19-Apr-1996 #text_change 21-Jul-2000
 C/Accession: S60266
 R/Greenberg, A.S.; Avila, D.; Hughes, M.; Hughes, A.; McKimney, E.C.; Flajnik, M.F.
 Nature 378, 168-173, 1995
 A/Title: A new antigen receptor gene family that undergoes rearrangement and extensive
 A/Reference number: S60266; MUID:95183140; PMID:7877689
 A/Accession: S60266
 A/Status: preliminary; nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-684 <GRE>
 A/Cross-references: EMBL:U18701; NID:g699442; PIDN:AA848195.1; PID:g699443

Query Match 9.6%; Score 326.5; DB 2; Length 684;

Best Local Similarity 23.8%; Pred. No. 1e-11;

Matches 156; Conservative 86; Mismatches 253; Indels 161; Gaps 29;

QY 16 LALLPAATQGNKRVLGKGDVTELTCTAS--QKSLQFMW-KNSNOIKILGNGSFLTKG 72
 Db 145 VSLHSATBEQA-----NRFQVLVCLISGYPENIAVSMQKNTKTI-----TSGFATIS 194
 QY 73 PSTLNDRAASRSL-----WDGQNFLLIKLKIEDSDTYICEVEDQKEVQLVFLG 125
 Db 195 PVTSSNDPSCASLLKVPLOEWSRGS-----VYSCQV----- 226
 QY 126 TANSDFHLQGSLLTLLSPSSPSVQCRSPRKNIGGKTLVSQLELDQSGTWCT 185

```

Db      227 -SH$T$N$Q$R$E$I$S$T-----S$E$A$V$L$L$R$D-----T$Y$E$I$W$I$D$S$A$T$L$V$C$E 268
Qy      186 V$L$O$N$Q$K$V$E$F$K$I$D$I$V$L$A$F$O$K$A$S$I$Y$Y$K$E$G$E$V$E---F$S$P$L$A$F$T$V$E$K$L$T$G$S$G$E$L$M$Q$A 242
Db      269 V$L$T$V$S$A$G-----V$V$S$M$W$M$N$G$K$V---R$B$G$V$E$M$P$E$T$K$M$G$N$G$Y$L$I$S$T$S$V$E$E$M--- 317
Qy      243 E$A$S$S$K$S$K$M$I$T$D$L$K$K$K$E$V$S$V$K$R$Y$T$O$D$P$K$L$O$M$G$K$U$P$H$L$T$P$O$A$L$P$O$A$G$S$N$L$T$A$E 302
Db      318 ----O$S$G$E$Y$T$S$A$K$D$O$S$T$P$V$K$K$T$R$A$R$A$E$P$T$K$-H$L$R$L$P$P$E$E$I$O$S$T$----- 366
Qy      303 A$K$T$G$K$L$H$G$E$V$N$L$V$M$R$A$T$O$L$K$V$L$T$G$E$W$G$P$T$S$P$K$L$M$L$S$L$K$E$N$K$E$A$K$S$K$E$K$P$W$V$L$N 362
Db      367 -----S$A$T$L$T$C$L$I$G$F$Y$D$K---V$S$V$M$Q$D$V$S$A$N-----V$T$N 399
Qy      363 P$E$A$G$M$Q$C$L$S$D$S$G$O$V$L$E$S$N$I$K$V$L$T$W$S$T---P$V$P$C$P$A$P$E$K$S-----C$D$K 406
Db      400 F$P$T$A$E$O$D$L$T$F$T$S$R$L$L---N$L$T$A$V---E$M$K$G$A$K$Y$T$C$H$A$S$H$P$S$Q$T$V$K$V$I$R$N$Q$K$Y$D$C$R$Q 455
Qy      407 T$H$T$C$E$L$L$G$P$S$V$F$L$P$P$R$K$P$O$L$M$S$R$P$E$V$Y$C$V$V$V$D$S$H$E$R$E---V$K$F$W$Y$D$G$Y$E$V$H$N 465
Db      456 T$D$I-----V$S$L$U$K$P---P$P$E$I$W$T$Q$O$A$T$T$V$C$E$V---Y$S$L$E$N$I$K$F$M$Q$N$G$V$E$R$K$K 503
Qy      466 A$K$T$E$R$E$Q$N$S$T$Y$R$V$S$V$L$T$V$L$H$O$M$L$N$G$E$K$Y$C$K$S$N$K$A$L$P$A$P$I$E$K$T$S$K$A$R---C$O$P$R$E 524
Db      504 G$Y$E$T$O$N$P$E$M$S$G$S$K$T$T$V$S$U$K$K$W$M$A$S$M$D$S$T$E$V$C$L$V$E$B$E$L$P$T$P$V$K$A$S$T$R$K$A$N$S$Q$M$P$R 563
Qy      555 P$O$V$Y$L$L$P$S$R$D$E$L---T$K$O$V$S$L$T$C$L$V$K$G$F$P$S$D$I$A$V$E$M$S$N$Q$---P$E$N$N$Y$K$T$P$V$L$D$S$G$S 582
Db      564 P$K$Y$U$L$L$P$S$T$D$E$I$D$E$N$S$A$T$L$M$C$L$A$N$F$H$P$A$E$I$Y$V$G$M$M$A$N$I$D$L$L$D$G$Y$R$Q$V$D$S$E$K$G$S 623
Qy      563 P$F$L$Y$S$K$L$T$V$D$K$R$M$O$G$N$F$C$S$G$V$M$E$A$L$H-----N$H$Y$T$O$S$L$S$L$P$S$O$L$D$E$T 631
Db      624 S$F$V$T$O$L$R$L$T$A$E$M$N$S$D$T$T$S$C$L$V$G$P$S$N$R$D$L$I$R$T$N$S$N$G$P$S$S$V$N$S$V$V$L$D$T 679

```

```

RESULT 79
C31933
Ig mu chain C region - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Mar-1990 #sequence_revision 30-Jun-1992 #text_change 20-Sep-1999
C:Accession: C31933
R:Schwager, J.; Mikoyak, C.A.; Steiner, L.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 2245-2249, 1988
A:Title: Amino acid sequence of heavy chain from Xenopus laevis IGM deduced from cDNA
A:Reference number: A94192; MUID:88176921; PMID:2451244
A:Accession: C31933
A:Molecule type: mRNA
A:Residues: 1-453 <SCH>
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match          9.5%  Score 324.5;  DB 2;  Length 453;
Beet Local Similarity 29.5%;  Pred. No. 8e-12;
Matches 98;  Conservative 54;  Mismatches 115;  Indels 65;  Gaps 14;

QY      342  LKLENKEKVKSRERKPV-----WYLNDEAGWMOCLL--SDSGOYLL 380
          : : : : :
DB      146  LKNGQGTGEGRAVEEPEDKKRGYEAATSYLSITKEMWLDI--LYSCVVEHAASGS-LQ 201
          : : : : :

QY      381  ESNIKVLPFTWSPVPCPAPEPKSCDKHTCCELLGSPSVFLPPPKKDTLMISRTPEVTC 440
          : : : : :
DB      202  EKNM-----SKSLMCDTPITPTSIQVITIP--PSL-----ESIEPKKSATLTC 242
          : : : : :

QY      441  VVVVDVSHEDPEVKFWMY--VDGVEVGNNAKTKRREEDYNS--TYRVVSVLTVLHODMLNCKE 497
          : : : : :
DB      243  LVSNMANSBDLRSLISWFKKSGTOEIPLKTLEIGDALYNNRRTSYKVGTTVACADSMNNDK- 301
          : : : : :

QY      498  YKCKSNKALPAPIEKTSKAKGQPREPOVYTLPPSRDELTKNQ--VSLTCLVKGFYPSDI 556
          : : : : :
DB      302  FVCKKTEHRLTSLMSEKVFLEFKENGVEYNTPSYVFPPPLELSKREIATITLCIVKGFSPEI 361
          : : : : :

QY      557  AVEW--ESNGQENNYKTP-----PVLDSDGSFFLYSKLTVDKSRWQQGNVSCSYMH 608
          : : : : :

```

Dd	362	FVFMHLHKNVPRKONTINTSINDELLPKGQSGKRFYLSLTIIDIKOMDAGDSFSCVVGH	4211
Oy	609	EALHNHYTKSLSLSPG-----LQLDETC	632
Dd	422	ESLPLQLTQRSIDKSGKRPNNVSLVSLDTC	453

RESULT 80

Ig upsilon chain - axolotl (fragment)
 C/Species: Ambystoma mexicanum (axolotl)
 C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Sep-1999
 C/Accession: S31436
 R/Fellah, J.S.; Miles, M.V.; Schwager, J.; Charlemagne, J.
 submitted to the EMBL Data Library, November 1992
 A/Description: cDNA sequence of Ambystoma mexicanum upsilon heavy Igy chain
 A/Reference number: S31436
 A/Accession: S31436
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-433 <FEEL>
 A/Cross-references: EMBL:X69492; NID:G62420; PID:G62421
 C/Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 9.4%; Score 320.5; DB 2; Length 433;

Matches 108; Conservative 65; Mismatches 162; Indels 111; Gaps 19;

216 G E O V E F S P L A T V E K L T G S G E L W Q A E R A S S S K S - - - - - W I T - - - - - F D L K N K E 260

34 GEPVEVSWKAGSTT---MPGKTFPAAQQQATALSTSSQIRIPAYEWGTSYSCSVKHKP 90

QY 261 VSV-----KRVTDPPKLOMGKKLPLHLTLPOALPOYAGSGNLTIALEAKTKLH 309

```

Db      91 TSTEIHKTITSAECKKATSPSVQ-----LQSSCADTDGNGSI----- 129

```

QY 310 QEVNLVVRATQLOKNLTCEVWGPTSPKLMLSLKENKEAKVSKREKPVVLNPEAGMWQ 369

Db 130 -----ELVCLISGYTPDNIQVRWL VNDKMAPIQGQTS-----PQK----- 165

Qy 370 CLSDSGVLESNIKVLPT-----PVPCPAPEPKSCDKTHTCPBLG--GPSVF 420

Db 166 ---DGQGTSTTSQINVTKS DWASGDKYTCKVEHPATSSRAEDTIHNCADSQTPYQPKVF 222

421 LPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPRE-----EQY 4755

Db 223 LIAPKARD-LYIANQPVICKITKMENS-D-SLSVTW-----KRREGPEEAAVISEQY 272

QY 476 ---NSTYRVSVLTVLHQDWLNCKEYKCKSVSNKALPAPIEKTISKAGQPREQVYTLPP 532

Db 273 IDSDGFTAMSYLNTKNEMERGDEFTCKVKHFDLPPLSRSVSKPTGRSFAPTMVFAF 332

533 SRDELTK-NQVSLTCLVKGFYPSDIAVEMESNGQ--PENNYKTTPLVLSDD---GSFEL 585

D6 333 HEMELANYDFVSLTCLVKSFSPPDIXIQMKQGSVIPSDKYVSMERQEQAGTAGLGIFYFS 392

QY 586 YSKLTVDKSRWQQGNVFSCSVMHEAL 611

DB 393 YSMUTIQKSDWDKRETFICVAHSAY 418

RESULT 81

Id alpha chain C region - rabbit (fragment)

C:\species: *Oryctolagus cuniculus* (domestic rabbit)
C:\Date: 28-Aug-1985 #sequence revision 28-Aug-1985 #text change 16-Jul-1999

C/accession: A021/4
R/knight, K.L.; Martens, C.L.; Stoklosa, C.M.; Schneiderman, R.D.

A; Title: Genes encoding alpha-heavy chains of rabbit IgA: characterization of

A;Accession: A02174

A:Molecule type: mRNA
A:Residues: 1-299 <KNI>
A:Cross-references: GB:X00353; NID:g1575; P1DN:CAA25100.1; PID:g1576
C:Comment: This immunoglobulin belongs to the Iga-g subclass. It was isolated from a rat
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)
chain disulfide bonds. In some cases, such as Iga and IgM, the subunits associate into 1a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin; plasma
F:86-152/Domain: immunoglobulin homology <IM1>
F:189-261/Domain: immunoglobulin homology <IM2>
F:138,286/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 306; DB 1; Length 299;
Best Local Similarity 32.2%; Pred. No. 1.7e-11;
Matches 88; Conservative 42; Mismatches 115; Indels 28; Gaps 13;

QY 369 QCLISDGGVLLSNIKVLPTWSTPVP--CPAPEKSCDKHTCP-----ELIG--PSV 419
DB 20 QCLGQKSAACHVEYNSVINESLPVPFDDCPA---NSC---CTCPSSSRNLISGQPSL 73
QY 420 FLPPPKXTLMSRPEVTCVVDVSHDEPVKFNWYDGVVHNAKTRPEEQYNSTY 479
DB 74 SLQRPDLGD-LLIGRDASLTCTLSGLKNPEDAV-FTW--EPTNGNEPVQORARDISGCY 129
QY 480 RIVSVLTVLHQMNGEKYCKVSNKALPA-PIEKTISKAKQPREPOVYTLPPSRDEL 538
DB 130 SVSSVLPSSAEWKARTEFTCTVTHPEIDSGSLTATISRGVTP--PQVHLPPPEELA 187
QY 539 KN-QVSLTCLVGFPSDIAVEMESNGQ--PENNY---KTPPVLDSDGSFPLYSKLTVD 592
DB 188 LNEQVTLTCLVGFSPKDVLSMRHQGEVPEDSFLVWKSMPSSODKATYATISLRVP 247
QY 593 KSRWQGNVFSQSVNHEALHNHYTKSLSLSPG 625
DB 248 AEDMNGDTYSCWVGHEGLAEHFTQRTIDRLAG 280

RESULT 82
S21461
T-cell surface glycoprotein CD4 (allele 1) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 21-Jan-2000
C:Accession: I47131; S21461
R:Guetafason, K.; Germana, S.; Sundt, T.M.
J. Immunol. 151, 1365-1370, 1993
A:Title: Extensive allelic polymorphism in the CDR2-like region of the miniature swine
A:Reference number: I47131; MUID:93329116; PMID:8335933
A:Accession: I47131
A:Status: preliminary; translated from GB/EMBL/DDAJ
A:Molecule type: mRNA
A:Residues: 1-99 <GU2>
A:Cross-references: EMBL:X65629; NID:g1928; P1DN:CAA46593.1; PID:g388232
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C:Keywords: glycoprotein; T-cell
F:3-81/Domain: immunoglobulin homology <IMM>

Query Match 8.9%; Score 305.5; DB 2; Length 99;
Best Local Similarity 60.2%; Pred. No. 1.5e-11;
Matches 59; Conservative 16; Mismatches 22; Indels 1; Gaps 1;

QY 32 KKSDTVLCTASQKSIQPFHMKNSQIKILGNQSFLL-TKGPSKINDRADRSRLMDQG 90
DB 1 KAGDIALPCHSSQKLPFNWKSNOTKILGHSFWHTASTELTSRLDSKKNWDHG 60

QY 91 NFFLIINKLKIEDSDTYICEVEDQKEVQLLYVCTLAN 128
DB 61 SFLIINKLEVTDSGYICEVEDKRIEVLQVLRFTAS 98

RESULT 83
S09276
Ig alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09276
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The Iga heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09264; MUID:90076124; PMID:2512120
A:Accession: S09276
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-338 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:228-300/Domain: immunoglobulin homology <IMM>

Query Match 8.9%; Score 303; DB 2; Length 338;
Best Local Similarity 31.9%; Pred. No. 1e-10;
Matches 87; Conservative 43; Mismatches 115; Indels 28; Gaps 13;

QY 369 QCLISDGGVLLSNIKVLPTWSTPVP--CPAPEKSCDKHTCP-----ELIG--PSV 419
DB 59 QCLGQKSAACHVEYNSVINESLPVPFDDCPA---NSC---CTCPSSSRNLISGQPSL 112
QY 420 FLPPPKXTLMSRPEVTCVVDVSHDEPVKFNWYDGVVHNAKTRPEEQYNSTY 479
DB 113 SLQRPDLGD-LLIGRDASLTCTLSGLKNPEDAV-FTW--EPTNGNEPVQORARDISGCY 168
QY 480 RIVSVLTVLHQMNGEKYCKVSNKALPA-PIEKTISKAKQPREPOVYTLPPSRDEL 538
DB 169 SVSSVLPSSAEWKARTEFTCTVTHPEIDSGSLTATISRGVTP--PQVHLPPPEELA 226
QY 539 KN-QVSLTCLVGFPSDIAVEMESNGQ--PENNY---KTPPVLDSDGSFPLYSKLTVD 592
DB 227 LNEQVTLTCLVGFSPKDVLSMRHQGEVPEDSFLVWKSMPSSODKATYATISLRVP 286
QY 593 KSRWQGNVFSQSVNHEALHNHYTKSLSLSPG 625
DB 287 AEDMNGDTYSCWVGHEGLAEHFTQRTIDRLAG 319

RESULT 84
HVKKCO
Ig mu chain C region, membrane-bound (clone 3050) - horn shark
C:Species: Heterodontus francisci (horn shark)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Aug-1996
C:Accession: S01854; C32716; A46530
R:Kokubu, F.; Hinder, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.
EMBO J. 7, 1979-1988, 1988
A:Title: Complete structure and organization of immunoglobulin heavy chain constant region
A:Reference number: S00980; MUID:88328985; PMID:3138109
A:Accession: S01854
A:Molecule type: DNA
A:Residues: 1-461 <KOK>
A:Cross-references: EMBL:X07781
A:Note: The sequence was determined from the germline gene
R:Kokubu, F.; Hinder, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.
Proc. Natl. Acad. Sci. U.S.A. 84, 5868-5872, 1987
A:Title: Extensive families of constant region genes in a phylogenetically primitive vertebrate
A:Reference number: A32716; MUID:87289703; PMID:3475706
A:Accession: C32716
A:Molecule type: DNA
A:Residues: 1-99 <K02>
A:Cross-references: GB:M17186
C:Genetics: 100/1; 206/1; 309/1; 419/1; 459/3
A:Introns: 100/1; 206/1; 309/1; 419/1; 459/3
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa)
chain disulfide bonds. In some cases, such as Iga and IgM, the subunits associate into 1a
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-87/Domain: immunoglobulin homology <IMM1>
F:123-190/Domain: immunoglobulin homology <IMM2>
F:228-291/Domain: immunoglobulin homology <IMM3>
F:330-400/Domain: immunoglobulin homology <IMM4>
F:438-458/Domain: transmembrane #status predicted <TMW>


```

Db      219 LGNN-EPVCAAGAAATCA-----DVKETTGGDVCVCFIFSKVTLTLLSDPQEDDERVLLV 272
Qy      283 TLPOALPOVAGSGNLTALAEKTKLHQEVNLVVMRATQIQKRLTCEWGPYSFKMLSL 342
Db      273 CLVEGLPVS-AGA-----AIQ-----W----- 287
Qy      343 KLENKAKYSKRRKPPVNLNPEAGMOCILSDSGOVLBSNIVL-PTW----- 390
Db      288 -LQDNBEMTPAPE-----SDSGCSDC--TESG-VTQMSRVVTKRSGEGAGQFCRYT 337
Qy      391 --STVPVCPAPEKSCDKHTCELLGSPVFLFPKPKDTLMISRTPEVTVVDSH 448
Db      338 HGLMKEVTVATVETDCAAT--POL-----QVSLPPTLEB-LIVSNATVTCVSNAAA 389
Qy      449 DPEVKENWYVD---GVEVNAKTKPREEQNSTYRVSVTLVHQMLNKEYCKVSNK 505
Db      390 D-GVSVWSRSSGSGGLDV-----SQTEDRQADGRVTVRSFLRVCAEEMNGGETFGCSVRE 444
Qy      506 ALPAPLEKTSKAKGP-REPOVYTLPPSRDELTKNOVSLTGLVGFYPSDIAVEMSN 563
Db      445 GV-VVAEESIRKKTDTPLHAPSVVFPPEBELSLQETATLTCTMASSFLPSSILLTWTOQ 503
Qy      564 GGP--ENNYKTTPVPLVSDSGFP-LYSKLTVDKSRMOQGNVFCSVHMEALHNHYTQSL 620
Db      504 NQGISPNVLIIFGP--EKQGDFFISYLSKLVSVEDMQRGDFGCVGVGHDSIPLNFIHKGI 561
Qy      621 SLSPG 625
Db      562 DKNAQ 566

```

RESULT 87

```

Ig alpha chain C region - rabbit (fragment)
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 29-Jan-1993 #sequence_rev150 29-Jan-1993 #text_change 16-Jul-1999
C/Accession: S09270
R/Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A/Title: The Iga heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A/Reference number: S09264; MUID:90076124; PMID:2512120
A/Accession: S09270
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-348 <BOR>
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin
F/133-199/Domain: immunoglobulin homology <IM>

```

```

Query Match      8.7%; Score 298.5; DB 2; Length 348;
Best Local Similarity 32.6%; Pred. No. 1.9e-10;
Matches 87; Conservative 41; Mismatches 114; Indels 25; Gaps 13;

Qy      369 QCLLSDSGOVLBSNIVLPTWSTVPVCPAPEKSCDKHT--TCPELLGSPVFLFPKPK 426
Db      78 QCLEYDSAAACHVEYNSVI--NESLPVPFPDP---CEQCCHPSCGE---PSLSLQRPDL 127
Qy      427 KDTLMISRTPEVTCVVVDVSHEDPEVKFNWY-VGVEVNAKTKPREEQNSTYRVSVTL 485
Db      128 RD-LLLGSDASLTCITRLGLKYPEDAV-FTWEPNTGNF--VQSPQRPDP--GCYVSSVYL 182
Qy      486 TVLHQMLNKEYCKVSNKALP-APLEKTSKAKGPREPOVYTLPPSRDELTKNQ-VS 543
Db      183 PGCAEPNNACTEFCVTHPEIRIGSLTATISKDTGSLTPVOYHLPPPEEELALNALVLT 242
Qy      544 LTCLVKGFPYSDIAVEMSNQ--OPENNY--KTPPVPLVSDSGFPLYSKLTVDKSRMOQ 598
Db      243 LTCLVKGFPYSDVLYVTKKQVDEVPENSLVWKLPLPSPGDDPTTYAVTSLIRVSAEDMWQ 302
Qy      599 GNVFSCSVHMEALHNHYTQSLSPG 625
Db      303 GDSYTCVVGHEGLAEHFTQRTIDREAG 329

```

RESULT 88

```

Ig alpha chain C region - pig (fragment)
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 21-Feb-1997 #sequence_rev150 21-Feb-1997 #text_change 21-Jan-2000
C/Accession: I47175
R/Brown, W.R.; Butler, J.E.
Mol. Immunol. 31, 633-642, 1994
A/Title: Characterization of a C alpha gene of swine.
A/Reference number: I47175; MUID:94254897; PMID:7545929
A/Accession: I47175
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-342 <BOR>
A/Cross-References: EMBL:U12594; NID:9555826; PIDN:AAA65943.1; PID:9555827
C/Genetics:
A/Gene: IGACalpha
C/Superfamily: immunoglobulin C region; immunoglobulin homology
F/232-304/Domain: immunoglobulin homology <IM>

```

```

Query Match      8.7%; Score 298; DB 2; Length 342;
Best Local Similarity 31.8%; Pred. No. 2e-10;
Matches 88; Conservative 42; Mismatches 109; Indels 38; Gaps 13;

```

```

Qy      372 LSDSGOVLBSNI--KVLPTWSTVPVCPAPEKSCDKHTCELLGSPVFLFPKPKDT 429
Db      88 LSKSSQSV---NVPCVLPVS---DPCP-----QCCK-----PSLSLQRPALAD- 124
Qy      430 LMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKTKPREEQNSTYRVSVTLVHL 489
Db      125 LILGSNASLITLSTGLKSE-GVSTWQPSGSK-DAVQASPRDSC-GCYVSSILPQGA 181
Qy      490 QDMLNGKEYCKVSNKALPAPLEKTSKAKGPREPOVYTLPPSRDELTKNQ-VSLTCLV 548
Db      182 DPMNGEFTSCAAHSELKSALTATITKPKVTPFPVOYHLPPPEEELALNELVTLCLV 241
Qy      549 KGFYPSDIAVEMSNQ--PENNY--KTPPVPLVSDSGFPLYSKLTVDKSRMOQGNVS 603
Db      242 RGFSPDVLVRLQGGQELPRDKLVWESLPPGQAIPFYAVTSVLRVDAEDWKQGDTPS 301
Qy      604 CSVMEALHNHYTQSLSPG----LQDERCAEAQ 636
Db      302 CMVGHFALPLATQKTIIDRLAKGPTHVNVSVMAEAE 338

```

RESULT 89

HVRK2

```

Ig mu chain C region (clone 12022) - horn shark (fragment)
C/Species: Heterodontus francisci (horn shark)
C/Date: 30-Jun-1991 #sequence_rev150 30-Jun-1991 #text_change 16-Jul-1999
C/Accession: S00980
R/Kokubu, F.; Hinds, K.; Litman, R.; Shambloot, M.J.; Litman, G.W.
EMBO J. 7, 1979-1988, 1988
A/Title: Complete structure and organization of immunoglobulin heavy chain constant reg1
A/Reference number: S00980; MUID:88528985; PMID:3138105
A/Accession: S00980
A/Molecule type: mRNA
A/Residues: 1-438 <KOK>
A/Cross-References: EMBL:X07784; NID:963963; PIDN:CAA30617.1; PID:963964
A/Note: The sequence was determined from the differentially expressed gene
C/Complex: An immunoglobulin heterodimeric subunit consists of two identical light (L)
chain disulfide bonds. In some cases, such as Iga and Igm, the subunits associate into 1-
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: duplication; glycoprotein; heterodimer; immunoglobulin
F/1-438/Domain: C region <CR>
F/20-87/Domain: immunoglobulin homology <IM>
F/223-190/Domain: immunoglobulin homology <IM>
F/228-291/Domain: immunoglobulin homology <IM>
F/330-400/Domain: immunoglobulin homology <IM>
F/166,200,245,275,374,411,415,425/Binding site: carbohydrate (Asn) (covalent) #status p

```

Query Match	8.7%;	Score 297.5;	DB 1;	Length 438;
Best Local Similarity	25.2%;	Pred. No. 3e-10;		
Matches 113;	Conservative 72;	Mismatches 167;	Indels 97;	Gaps 19

```

Qy 208 SSIYKKEGEQVEFEPLAFVY-----EKLGSGBELWQABRASSKSM 251
Db 37 TSIIWKKCKDEPTTGLKITYSVLANKKGYTRSSOLTTSESVGSSKIYCEVRGES--LW 94
Qy 252 ITFDLKNKEVSKYRVTODPKLQMGAKLPLHLPLPOLPQAGSNDLTALAEAKTKHOE 311
Db 95 I-----KEI-----LDCK---GDIVPPTVILTQSSSEISRFRPATVLCGIIIDPHRS 139
Qy 312 VNLVYMRATQ-----LQKULTCEWNGPSPKMLSLKLENKAAYSKREKPVWLNPBG 366
Db 140 ITVSMLKKGOPMDSGFVTSPTCEVNGNSATSRLTVPAGE-----WFSNT--- 184
Qy 367 MMOCILSDSGQVLTESINIKVLPETMSVPVCAPEAPKSCDKHTICEPLLGGPVLPLPKP 426
Db 185 VYTOQVARG-ETTGSRNIT-----GSQVPC-----IDPVIKLLPBI 222
Qy 427 KOTLMISTRPEYTCVVDVSHEDPEYKENVYDVGEVHNAKTKPBEQY----NSTYRV 481
Db 223 EQVL-LEATVTLTCV---VSNAPYGVNSW-----TOEKPLKSEIAVOPGEDSDSV 270
Qy 482 VSVLTLVHODMNGEYKCKVKNKALPPIEKTISKAGOP-REPOV-YLPPSRDELTX 539
Db 271 ISTVWISTQAMLSGAEFYCVVSHOULPTPLRASIKKEEKVDUREPVSYLLEPAEDVASQ 330
Qy 540 NOVSLTCLVKGFPSPDIAMWEMSNGOPEN--NYKTTPLVSDG--SFLYSKLTYDKSN 596
Db 331 RFLSLTCLVKGSPSPREIFIKMTVNNKSVNPGNYKOTVEAENDRNSFFIYSLLSIAEBW 390
Qy 597 QOGANVFGSCWHEALHNHYTKSLSLSLSPG 625
Db 391 ASGASVSCVCGHEALPLKIINRTVKKSSG 419

```

RESULT 90
S09269
Ig alpha chain C region - rabbit (fragment)
C/Species: Oryctolagus cuniculus (domestic rabbit)
C/Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C/Accession: S09269
R/Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A/Title: The 15A heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A/Reference number: S09264; M01D:90076124; PMID:2512120
A/Accession: S09269
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-357 <BUR>
C/Superfamily: immunoglobulin C region; immunoglobulin homology
C/Keywords: immunoglobulin
E/142-208/Domain: immunoglobulin homology <IMM>

Query Match	8.7%	Score 297	DB 2	Length 357
Best Local Similarity	32.9%	Pred. No. 2,4e-10		
Matches	82	Conservative	38	Mismatches 99, Indels 30, Gaps 11

QY	388	PTWSPVYVCPAPBPSPSCDKHTTCPELGLSPSYFLFPPKPKDTLMTSRPEVTCVVVDYD	447
Db	109	PSDITTCPCPCSP-SC-----GEPISLIRPFLRD--LLTNSNASTLCTTGLRGLN	156
QY	448	EDPEYKFMVYDGVENHAKTRREEQYNS---TYRVSVLTVLHODWLNGEKYCKV	503
Db	157	PEGAV-FPM-----BPTNG-NKRPVQGSVQSYCGCGSVSVLPGCALPEPNACTEFTCTVT	209
QY	504	NKALP-APLEKTSISKAGQPREPOVYTLPPSDELTKNQ-VSLTCLVYGVSFSDYLAWE	561
Db	210	HPEIGGGLTAKISDTGAIIPVPHLPPEBELALNELVLTCLVKGFSKDYLVYT	269
QY	562	SNG--QPENNY--KTTPEVLVDSGSFPLYSKLTIVDKSRMOGNAVFSGVNHEALHNHYT	616

Db 270 NKGWNPENSLYMKFLPFGQEPITYAVTSLRRVPAEDNNQNESYTCVGHESLAHFT 322

QY 617 QKSLSLSPG 625
|:::|
Db 330 QRTIDRLSG 338

```

RESULT 91
S09274
Ig alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09274
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IGA heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference numbers: S09284; MUID:90076124; PMID:2512120
A:Accession: S09274
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-147 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
P:132-198/Domain: immunoglobulin homology <IMM>

```

Query Match	8.7%	Score 295.5	DB 2	Length 347
Best Local Similarity	33.2%	Pred. No. 2.9e-10		
Matches	89	Conservative 40	Mismatches 112	Indels 27
			Gaps	14
Qy	369	QCLTSDSGQVLLEENIKVLPWTSTPVCPEAPEKSCDKTH--TCPELIGSPVFLPPRP	426	
Db	77	QCLEYDGAACHVEINSVI--NESLPVPPPD----CEGCHPSCBE---PLSLSQRDL	126	
Qy	427	KDTLMISRTPEVTCVVVDVSHEDDE-VKFN-YYDGVENHNAKTRPREQYNSTYRVVSV	484	
Db	127	RD-LLLGSDASLTCTLGL--KDEGAIVFTGPTNGNE--PVQDSQRDPC-GCYSAVS	180	
Qy	485	LTVLHQWLNGKEYCKKCVKSNKALP-APLEKTIISAKGQPRRPQYTLTPPSDELTKNQ-V	542	
Db	181	LPGCAEPMNAGTEFECTVTHPELIGSSSLTATISKDTGSLTPPLVHLTPPSEELANLTV	240	
Qy	543	SLTCLVNGFYPSDLAVEMESNG--QPEENY---KTPPVLDSDGSFYLKLTLYDKSRNQ	597	
Db	241	TLTCLVNGFSKPDVLVSMTNKGVPKPSFLWVKRLPEPGQDPTTAVTSLLRPAEDWN	300	
Qy	598	QGNVSCSVNHEALHNHYTKSLSPG	625	
Db	301	QNESYSCVVAHEGLAEHFTQRTIIRLNG	328	

RESULT 92
B22360
Ig alpha-2 chain C region (allotype A2m(1)) - human
C:Species: Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C:Accession: B22360
R:Planagan, J.G.; Lefranc, M.P.; Rabbits, T.H.
Cell 36, 681-688, 1984
A:Title: Mechanisms of divergence and convergence of the human immunoglobulin alpha-1
A:Reference number: A94653; MUID:84110179; PMID:6621489
A:Accession: B22360
A:Molecule type: DNA
A:Residues: 1-340 <FLA>
C:Genetics:
A:Gene: GDB:IGHA2
A:Cross-references: GDB:119333; OMIM:147000
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1 103/1: 210/1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
;230-302/Domain: immunoglobulin homology <IM>

Query Match 8.6%; Score 293.5; DB 2; Length 340;

RESULT 95
S09266

```

C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09266
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IgH heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09264; MUID:90076124; PMID:2512120
A:Accession: S09266
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-352 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
P:241-314/Domain: immunoglobulin homology <IMM>

Query Match      8.4%; Score 285.5; DB 2; Length 352;
Beet Local Similarity 33.1%; Pred. No. 1.1e-09;
Matches 89; Conservative 36; Mismatches 101; Indels 43; Gaps 16;

Oy 374 DSGVLLNESNIKVL-----PTWSTPVCPPADPEKSCDKTHTCPPELLGSPVFLPPKPK 427
Db 91 DEGG-----NLTVLPECKDPN-SDPTPCPCP-PIYC-----GEPSTLQPPDIG 133
Oy 428 DTLMISRPPEVYCVAVDSHEDPE-VKENWY-VDGVEVNAATKPREDOYN-STRYVSV 484
Db 134 D-LILIESNASLTCITSLGL--KDPEGAVFTWNPFTNGEYVQGST---QSPYCGCYSSSV 186
Oy 485 LTVLHQDWLNGKEVYCKVSNKALP-APIEKTSKAKGQPREQVYTLTPSRDELTKNQ-V 542
Db 187 LPGCAEPNNAAGIEFTCTVHPRIEGLSLATISRSGLTP--PQVHLPPTEELALNALV 244
Oy 543 SLTCLVKGFPSPDIADVWESNQ--PENNY--KTPPEVLDSGSPFLYSKLTVDKSRW 596
Db 245 TLTCVLRGSPKPDVLSWTHNGTLVVPKDSFLWPKRLPPEGEPITYAVTSLRLVPAEDW 304
Oy 597 QCGNIFSGSVNHEALHNHTQKSLSPG 625
Db 305 NQGDSDYSCVGHGELAHEFTORTIDRLAG 333

RESULT 96
S09265
I9 alpha chain C region - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S09265
R:Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 8, 4041-4047, 1989
A:Title: The IgH heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A:Reference number: S09264; MUID:90076124; PMID:2512120
A:Accession: S09265
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-357 <BUR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
P:246-319/Domain: immunoglobulin homology <IMM>

Query Match      8.4%; Score 285.5; DB 2; Length 357;
Beet Local Similarity 32.4%; Pred. No. 1.2e-09;
Matches 83; Conservative 38; Mismatches 104; Indels 31; Gaps 10;

Oy 381 ESNIVLPLTWSTPVCPPADPEKSCDKTHTCPPELLGSPVFLPPKPKDTLMISRPPEVTC 440
Db 103 ESTIEPPTPTPCPCPSPD--SC-----GKPSISLQRPDLGD-LILINSNASLTC 148
Oy 441 VVVDVSHEDPEVYKFMVYDGVENHNAKTPREE--QYNSTRVYVSVLTVLHQDWLNGKE 497
Db 149 TLRLGLNPEGAV-FTW-----EPTFGKEPVGQSPQLDHCQCYSSSVLPGCAVLMNAGTE 202
Oy 498 YKCKRSNKAALP-PLEKITSKAKGQPREQVYTLTPSRDELTKNQ-VSLTCLVKGFPSPD 555
Db 203 FCTCTVTHPEIREDSDLTGTSIGTSLIPQVHLPLPPSEELALNALVTLTCVLRGFSPPD 262

```

```

Oy 556 IAVENESNGQ---PENNKKTPPEPLD---SDGSFPLYSLTLYDSCRMOQGVFSCSVHME 609
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 263 VLVSTHNGTTPVPEKOSTLVKPKLPKPGDPPTTITSLRVPADMDNGSDSYCVHGE 322
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 610 ALAHNYTOKSLSLSPG 625
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 GLAEHFTOKTIDRLAG 338
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 97
S09272
Ig alpha chain C region - rabbit (fragment)
C.Species: Oryctolagus cuniculus (domestic rabbit)
C.Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C.Accession: S09272
R.Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.
EMBO J. 9, 4041-4047, 1989
A.Title: The Ig heavy-chain gene family in rabbit: cloning and sequence analysis of 13
A.Reference number: S09264; MUID:90076124; PMID:2512120
A.Accession: S09272
A.Status: not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-343 <BUR>
C.Superfamily: immunoglobulin C region; immunoglobulin homology
C.Keywords: immunoglobulin
F.232-305/Domain: immunoglobulin homology <IM>

Query Match 8.3%; Score 285; DB 2; Length 343;
Best Local Similarity 32.1%; Pred. No. 1.2e-09;
Matches 89; Conservative 37; Mismatches 111; Indels 40; Gaps 14;

Oy 363 PEAGMOCQLS---DSGOVLBSNLIKVLPTMSTPVPCEAPEKSCDKTHTCCELLGSPV 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 74 PRDGAVTCHVHNVDGQDL-----TVPCCDCHCYCP---PTSC-----GEPGL 114
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 420 FLPPPKPQDTIMISRTPEVTCVVVDVSHEDPE-VKENNY-VDGVEVHNAKTPREEQYN- 476
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 115 SLQRPDIDDLLEBSA-SLTCTLSGL--KDEGAVFTPEPTNG---NEFVQSVQSYPC 167
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 477 STYRVSVLYVLHODMNGKEYKCKVSNKALP-ADIEKTSIKAKQPREPOVYTLPPSHD 535
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 GCYSVSVSLPCGAEPNMGAGTEFTCTVTHPEIGSGLTATISISRSGLTPPVHLLPPPE 227
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 536 ELTKK-QVSLCIKVGKFPSPDIIVAVESNGQ---PENN---KTPPVLDSDGSFPLYSK 588
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 228 ELALNEQVTLTCLVGFSPKVDLVSWTNGTLVVPKDSYLWVKPLPEPGDPPTTAAVYSL 287
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 589 LTVDSKRWQGNVFCSCVHAEALAHNYTOKSLSLSPG 625
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 288 LRVSADMDNGSDSYCVHGEGLAEHFTORTIDRQAG 324
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 98
I56230
Ig alpha-2 chain - human (fragment)
C.Species: Homo sapiens (man)
C.Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000
C.Accession: I56230
R.Chintalacharuvu, K.R.; Ratnes, M.; Morrison, S.L.
J. Immunol. 152, 5299-5304, 1994
A.Title: Divergence of human alpha-chain constant region gene sequences. A novel recomb
A.Reference number: I56230; MUID:94246170; PMID:8189047
A.Accession: I56230
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-340 <RES>
A.Cross-references: GB:S71043; NID:G546798; PIDN:AMB30803.1; PID:G546799
C.Genetics:
A.Gene: IGA2
A.Introns: 103/1; 210/1
C.Superfamily: immunoglobulin C region; immunoglobulin homology
F.230-302/Domain: immunoglobulin homology <IM>

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 3, 2004, 13:14:00 ; Search time 54.8475 Seconds
(without alignments)
3706.029 Million cell updates/sec

Title: SEQ7
Perfect score: 3414
Sequence: 1 MNRGVPRHLLVLQLALP.....DETCAEADGELDGLMTTDP 648

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2116	62.0	530	8 US-08-485-163-5	Sequence 5, Appl1
2	2116	62.0	530	9 US-09-766-995-4	Sequence 4, Appl1
3	2059	60.3	432	8 US-08-485-163-3	Sequence 3, Appl1
4	2059	60.3	432	9 US-09-766-995-2	Sequence 2, Appl1
5	2036.5	59.7	532	10 US-09-939-537-6	Sequence 6, Appl1
6	2036.5	59.7	532	11 US-09-243-008-6	Sequence 6, Appl1
7	2036.5	59.7	575	10 US-09-939-537-4	Sequence 4, Appl1
8	2036.5	59.7	575	11 US-09-243-008-4	Sequence 4, Appl1
9	2032.5	59.5	462	10 US-09-939-537-5	Sequence 5, Appl1
10	2032.5	59.5	462	11 US-09-243-008-5	Sequence 5, Appl1
11	2029	59.4	398	10 US-09-939-537-29	Sequence 29, Appl1
12	2021	59.2	457	11 US-09-891-119A-9	Sequence 9, Appl1
13	2017	59.1	402	14 US-10-097-044A-1	Sequence 1, Appl1
14	2015	59.0	458	14 US-10-103-597A-39	Sequence 39, Appl1
15	2015	59.0	458	14 US-10-188-44A-39	Sequence 39, Appl1

16	2015	59.0	458	14 US-10-207-655-170	Sequence 170, App
17	2007	58.8	458	12 US-10-151-274-3	Sequence 3, Appl1
18	2001	58.6	397	11 US-09-891-119A-2	Sequence 2, Appl1
19	1997	58.5	458	8 US-08-681-219-27	Sequence 27, Appl1
20	1997	58.5	458	11 US-09-230-111C-25	Sequence 25, Appl1
21	1997	58.5	458	14 US-10-092-138-25	Sequence 25, Appl1
22	1904	55.8	434	14 US-10-097-044A-4	Sequence 4, Appl1
23	1891	55.4	370	9 US-09-759-841-6	Sequence 6, Appl1
24	1877.5	55.0	448	14 US-10-024-329-32	Sequence 32, Appl1
25	1877.5	55.0	448	14 US-09-939-537-33	Sequence 33, Appl1
26	1877.5	55.0	448	14 US-10-363-427-18	Sequence 18, Appl1
27	1877.5	55.0	448	14 US-10-363-427-22	Sequence 22, Appl1
28	1877.5	55.0	448	14 US-10-363-427-14	Sequence 14, Appl1
29	1877.5	55.0	448	14 US-09-935-868-8	Sequence 8, Appl1
30	1877.5	55.0	448	14 US-10-287-035-8	Sequence 8, Appl1
31	1877.5	55.0	448	14 US-10-282-162-8	Sequence 8, Appl1
32	1877.5	55.0	448	14 US-09-815-108-22	Sequence 22, Appl1
33	1877.5	55.0	448	14 US-10-229-584-22	Sequence 22, Appl1
34	1877.5	55.0	448	14 US-09-910-600-16	Sequence 16, Appl1
35	1877.5	55.0	448	14 US-10-207-655-348	Sequence 348, App
36	1877.5	55.0	448	14 US-10-207-655-345	Sequence 345, App
37	1877.5	55.0	448	14 US-10-207-655-344	Sequence 344, App
38	1877.5	55.0	448	14 US-10-207-655-344	Sequence 344, App
39	1877.5	55.0	448	14 US-09-773-877A-12	Sequence 12, Appl1
40	1877.5	55.0	448	14 US-09-773-877A-20	Sequence 20, Appl1
41	1877.5	55.0	448	14 US-10-412-406-33	Sequence 33, Appl1
42	1877.5	55.0	448	14 US-10-207-655-266	Sequence 266, App
43	1877.5	55.0	448	14 US-10-412-406-32	Sequence 32, Appl1
44	1877.5	55.0	448	14 US-10-207-655-350	Sequence 350, App
45	1877.5	55.0	448	14 US-09-875-338-9	Sequence 9, Appl1
46	1877.5	55.0	448	14 US-10-077-023-9	Sequence 9, Appl1
47	1877.5	55.0	448	14 US-10-077-023-15	Sequence 15, Appl1
48	1877.5	55.0	448	14 US-10-207-655-240	Sequence 240, App
49	1877.5	55.0	448	14 US-10-207-655-398	Sequence 398, App
50	1877.5	55.0	448	14 US-10-683-255-2	Sequence 2, Appl1
51	1877.5	55.0	448	14 US-10-232-838-18	Sequence 18, Appl1
52	1877.5	55.0	448	14 US-09-875-338-5	Sequence 5, Appl1
53	1877.5	55.0	448	14 US-10-077-023-5	Sequence 5, Appl1
54	1877.5	55.0	448	14 US-10-207-655-15	Sequence 15, Appl1
55	1877.5	55.0	448	14 US-10-207-655-148	Sequence 148, App
56	1877.5	55.0	448	14 US-10-053-530-15	Sequence 15, Appl1
57	1877.5	55.0	448	14 US-10-416-011-2	Sequence 2, Appl1
58	1877.5	55.0	448	14 US-09-875-338-17	Sequence 17, Appl1
59	1877.5	55.0	448	14 US-10-077-023-17	Sequence 17, Appl1
60	1877.5	55.0	448	14 US-10-679-620-64	Sequence 64, Appl1
61	1877.5	55.0	448	14 US-10-463-260-6	Sequence 6, Appl1
62	1877.5	55.0	448	14 US-10-120-198B-2	Sequence 2, Appl1
63	1877.5	55.0	448	14 US-10-163-427-24	Sequence 24, Appl1
64	1877.5	55.0	448	14 US-10-404-724-8	Sequence 8, Appl1
65	1877.5	55.0	448	14 US-10-232-838-17	Sequence 17, Appl1
66	1877.5	55.0	448	14 US-10-357-653-2	Sequence 2, Appl1
67	1877.5	55.0	448	14 US-10-282-162-52	Sequence 52, Appl1
68	1877.5	55.0	448	14 US-10-679-620-62	Sequence 62, Appl1
69	1877.5	55.0	448	14 US-10-683-255-6	Sequence 6, Appl1
70	1877.5	55.0	448	14 US-10-683-255-4	Sequence 4, Appl1
71	1877.5	55.0	448	14 US-10-207-655-397	Sequence 397, App
72	1877.5	55.0	448	14 US-10-452-646-9	Sequence 9, Appl1
73	1877.5	55.0	448	14 US-10-207-655-396	Sequence 396, App
74	1877.5	55.0	448	14 US-10-163-427-20	Sequence 20, Appl1
75	1877.5	55.0	448	14 US-10-207-655-346	Sequence 346, App
76	1877.5	55.0	448	14 US-10-107-991B-3	Sequence 3, Appl1
77	1877.5	55.0	448	14 US-10-363-427-16	Sequence 16, Appl1
78	1877.5	55.0	448	14 US-10-334-235-38	Sequence 38, Appl1
79	1877.5	55.0	448	14 US-10-418-836-39	Sequence 39, Appl1
80	1877.5	55.0	448	14 US-10-207-655-270	Sequence 270, App
81	1877.5	55.0	448	14 US-10-418-836-38	Sequence 38, Appl1
82	1877.5	55.0	448	14 US-10-232-838-15	Sequence 15, Appl1
83	1877.5	55.0	448	14 US-10-232-838-15	Sequence 15, Appl1
84	1877.5	55.0	448	14 US-10-207-655-307	Sequence 307, App
85	1877.5	55.0	448	14 US-09-773-877A-18	Sequence 18, Appl1
86	1877.5	55.0	448	14 US-10-433-108-29	Sequence 29, Appl1
87	1877.5	55.0	448	14 US-10-404-724-23	Sequence 23, Appl1
88	1877.5	55.0	448	14 US-10-404-724-23	Sequence 23, Appl1


```

89 1228 36.0 467 16 US-10-656-769-32
90 1228 36.0 476 9 US-09-747-669-3
91 1228 36.0 476 14 US-10-290-703-3
92 1227.5 36.0 472 12 US-10-433-108-21
93 1227.5 36.0 448 15 US-10-378-567-2
94 1226 35.9 471 15 US-10-108-260A-4285
95 1225.5 35.9 470 15 US-10-108-260A-4292
96 1225.5 35.9 900 15 US-10-282-162-34
97 1225 35.9 488 12 US-10-683-255-12
98 1224.5 35.9 366 16 US-10-193-616-14
99 1224.5 35.9 541 16 US-10-471-151-32
100 1224.5 35.9 558 16 US-10-471-151-31
101 1224 35.9 500 14 US-10-207-655-274
102 1223.5 35.8 280 12 US-10-433-108-27
103 1223.5 35.8 470 14 US-10-216-484-117
104 1223.5 35.8 470 14 US-10-216-484-113
105 1223.5 35.8 470 14 US-10-384-933-117
106 1223.5 35.8 470 14 US-10-384-933-113
107 1223.5 35.8 499 14 US-10-207-655-17
108 1223.5 35.8 499 14 US-10-053-530-17
109 1223 35.8 455 12 US-10-404-724-25
110 1223 35.8 480 14 US-10-077-023-133
111 1223 35.8 480 14 US-10-077-023-135
112 1222.5 35.8 445 14 US-10-320-231A-79
113 1222.5 35.8 470 14 US-10-216-484-147
114 1222.5 35.8 470 14 US-10-384-933-147
115 1222.5 35.8 475 9 US-09-740-002-25
116 1222.5 35.8 475 16 US-10-325-698-25
117 1222.5 35.8 579 14 US-10-138-727A-41
118 1222.5 35.8 680 8 US-08-469-583A-15
119 1222 35.8 915 14 US-10-282-162-46
120 1221.5 35.8 444 14 US-10-150-475A-6
121 1221.5 35.8 444 16 US-10-704-522-6
122 1221.5 35.8 444 16 US-10-645-215-6
123 1221.5 35.8 470 14 US-10-216-484-145
124 1221.5 35.8 470 14 US-10-384-933-145
125 1221 35.8 474 15 US-10-108-260A-4282

```

ALIGNMENTS

```

Sequence 32, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 21, Appl
Sequence 2, Appl
Sequence 4285, Appl
Sequence 4292, Ap
Sequence 34, Appl
Sequence 12, Appl
Sequence 14, Appl
Sequence 32, Appl
Sequence 31, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 117, App
Sequence 117, App
Sequence 117, App
Sequence 143, App
Sequence 17, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 135, App
Sequence 79, Appl
Sequence 147, App
Sequence 147, App
Sequence 25, Appl
Sequence 25, Appl
Sequence 41, Appl
Sequence 15, Appl
Sequence 46, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 6, Appl
Sequence 145, App
Sequence 145, App
Sequence 4282, Ap

```

```

RESULT 1
; Sequence 5, Application US/08485163
; Publication No. US2002098191A1
; GENERAL INFORMATION:
; APPLICANT: Beaudry, Gary A.
; APPLICANT: Maddon, Paul J.
; TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,163
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 37690-II-1-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400

```

```

; TELEFAX: (212) 391-0525
;
;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
;
; MOLECULE TYPE: CDNA
;
; ORIGINAL SOURCE:
; ORGANISM: homo sapien
; CELL TYPE: lymphocyte
;
; US-08-485-163-5
;
; Query Match 62.0%; Score 2116; DB 8; Length 530;
; Best Local Similarity 68.9%; Pred. No. 1.7e-141;
; Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;
;
; 1 MKRGVPRHLILVLTALIPATQGNKVLAGKGPVETLTCTASQKSIQFMKNSNQIK 60
; 1 MKRGVPRHLILVLTALIPATQGNKVLAGKGPVETLTCTASQKSIQFMKNSNQIK 60
; 1 MKRGVPRHLILVLTALIPATQGNKVLAGKGPVETLTCTASQKSIQFMKNSNQIK 60
; 61 ILGNQSPFLTKGPSKLNDRADSRSLMDQGNPILIKLIKIEDPTTYICEVEDQKEVOL 120
; 61 ILGNQSPFLTKGPSKLNDRADSRSLMDQGNPILIKLIKIEDPTTYICEVEDQKEVOL 120
; 61 ILGNQSPFLTKGPSKLNDRADSRSLMDQGNPILIKLIKIEDPTTYICEVEDQKEVOL 120
; 121 LVFGILTANDTLLDQGSITLTLSPSGSPVQCRSPGKNIQGGKTLVSQLELDQSG 180
; 121 LVFGILTANDTLLDQGSITLTLSPSGSPVQCRSPGKNIQGGKTLVSQLELDQSG 180
; 121 LVFGILTANDTLLDQGSITLTLSPSGSPVQCRSPGKNIQGGKTLVSQLELDQSG 180
; 181 TWTCVTLQKQKVEFKIDIVLAFOKASIVKKEGEQVEFSPFLATVEKLTGSGELMW 240
; 181 TWTCVTLQKQKVEFKIDIVLAFOKASIVKKEGEQVEFSPFLATVEKLTGSGELMW 240
; 181 TWTCVTLQKQKVEFKIDIVLAFOKASIVKKEGEQVEFSPFLATVEKLTGSGELMW 240
; 241 QERASSKSKWITFDLKNKEVSVKRVTDPKLQMGKULPLHLTLPLQALPOYAG--SGNL 297
; 217 -----PCSRSTSESTAALGCLVDYPPPEVTVSMNSGALTSGVH 255
;
; 298 TLALEAKTKLHAEVNLVVMRATOL-QKULTEVWCPSPKMLSLKLNKAKSKRRK 356
; 256 TEPAYLQSSGLYSLSVTVTPSSNFGOTQTYTCNV-----DHK 292
; 357 PVWVLPNPEAGMWQCLSDSGQVLLESNIKVLPTWSTPV-----PCPAPRPSCKDHTHP 411
; 293 P-----SNTKVDITVERKCCVECPRCARP----- 317
;
; 412 ELLEGDSVFLPPPKPKDTLMISRTPEVTCVVDVSHEDPEVKNMYVDGVEYNAATKPR 471
; 412 ELLEGDSVFLPPPKPKDTLMISRTPEVTCVVDVSHEDPEVKNMYVDGVEYNAATKPR 471
; 412 ELLEGDSVFLPPPKPKDTLMISRTPEVTCVVDVSHEDPEVKNMYVDGVEYNAATKPR 471
; 318 --VAGPSVFLPPPKPKDTLMISRTPEVTCVVDVSHEDPEVQFNWYVDGVEYNAATKPR 375
;
; 472 EQQYNSTVAVSVLTALHODMNLNGEKYKCKVKNKALPAIEKTISSAKGQPREPOVYTLR 531
; 472 EQQYNSTVAVSVLTALHODMNLNGEKYKCKVKNKALPAIEKTISSAKGQPREPOVYTLR 531
; 472 EQQYNSTVAVSVLTALHODMNLNGEKYKCKVKNKALPAIEKTISSAKGQPREPOVYTLR 531
; 376 EQQYNSTVAVSVLTALHODMNLNGEKYKCKVKNKALPAIEKTISSAKGQPREPOVYTLR 435
;
; 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTIPVLDSDGSFFLYSKLTV 591
; 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTIPVLDSDGSFFLYSKLTV 591
; 532 PSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTIPVLDSDGSFFLYSKLTV 591
; 436 PSREKTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTIPVLDSDGSFFLYSKLTV 495
;
; 592 DKSRLQGNVVFSCSVMEALAHNYTQKSLSPG 625
; 496 DKSRLQGNVVFSCSVMEALAHNYTQKSLSPG 529

```

```

RESULT 2
; Sequence 4, Application US/09766995
; Patent No. US20020052481A1
; GENERAL INFORMATION:
; APPLICANT: Graham P. Allaway et al.
; TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED CD4-GAMMA2 AND CD4-IgG2 IMMUNOCON
; FILE REFERENCE: 2048/41215-CB/JPM/SHS
; CURRENT APPLICATION NUMBER: US/09/766,995

```

/ CURRENT FILING DATE: 2001-01-22
 / NUMBER OF SEQ ID NOS: 9
 / SOFTWARE: Patentin version 3.0
 / SEQ ID NO 4
 / LENGTH: 530
 / TYPE: PRT
 / ORGANISM: homo sapiens
 / US-09-766-995-4

Query Match 62.0%; Score 2116; DB 9; Length 530;
 Best Local Similarity 68.9%; Pred. No. 1.7e-141;
 Matches 437; Conservative 25; Mismatches 58; Indels 114; Gaps 10;

QY 1 MNRGVPFRHLLVQLALPAAATQGNKVLGKGGDTVELTCTASOKKSIQFHMNSNQIK 60
 DB 1 MNRGVPFRHLLVQLALPAAATQGNKVLGKGGDTVELTCTASOKKSIQFHMNSNQIK 60
 QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
 DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
 QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180
 DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180
 QY 181 TWCTCTVLQNGKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
 DB 181 TWCTCTVLQNGKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
 QY 241 QAEBAASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPULHLPQALPOYAG---SGNL 297
 DB 241 QAEBAASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPULHLPQALPOYAG---SGNL 297
 QY 298 TLAEAKTGKLGHOENVLVMMARATQLOKNLICEVWGPTSPKMLSLKENKAIVSKREK 356
 DB 298 TLAEAKTGKLGHOENVLVMMARATQLOKNLICEVWGPTSPKMLSLKENKAIVSKREK 356
 QY 357 PAVWLINPEAGMOCCLSDSGVLLSNIKVLPTWSTPV-----PCPAPPEKCDKTHCP 411
 DB 357 PAVWLINPEAGMOCCLSDSGVLLSNIKVLPTWSTPV-----PCPAPPEKCDKTHCP 411
 QY 412 ELGQPSVFLPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPR 471
 DB 412 ELGQPSVFLPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPR 471
 QY 318 --VAGPSVFLPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPR 375
 DB 318 --VAGPSVFLPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPR 375
 QY 472 EEOQNSTFRVSVTLVHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLR 531
 DB 472 EEOQNSTFRVSVTLVHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLR 531
 QY 376 EEOQNSTFRVSVTLVHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLR 435
 DB 376 EEOQNSTFRVSVTLVHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLR 435
 QY 532 PSRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTV 591
 DB 532 PSRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTV 591
 QY 436 PSRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTV 495
 DB 436 PSRDELTKNOVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTV 495
 QY 592 DKSRMOQGNVPSCSVMHEALHNHYTOKSLSPG 625
 DB 592 DKSRMOQGNVPSCSVMHEALHNHYTOKSLSPG 625
 QY 496 DKSRMOQGNVPSCSVMHEALHNHYTOKSLSPG 529
 DB 496 DKSRMOQGNVPSCSVMHEALHNHYTOKSLSPG 529

RESULT 3
 US-08-485-163-3
 / Sequence 3, Application US/08485163
 / Publication No. US20020098191A1
 / GENERAL INFORMATION:
 / APPLICANT: Beaudry, Gary A.
 / APPLICANT: Maddon, Paul J.
 / TITLE OF INVENTION: CD4-GAMMA2 CD4-IgG2 CHIMERAS
 / NUMBER OF SEQUENCES: 10
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Cooper & Dunham LLP
 / STREET: 1185 Avenue of the Americas
 / CITY: New York
 / STATE: New York
 / COUNTRY: USA

ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.24
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,163
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 37690-II-1-PCT-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 TELEX:
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 432 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: homo sapien
 CELL TYPE: lymphocyte
 US-08-485-163-3

Query Match 60.3%; Score 2059; DB 8; Length 432;
 Best Local Similarity 65.6%; Pred. No. 1.4e-137;
 Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVPFRHLLVQLALPAAATQGNKVLGKGGDTVELTCTASOKKSIQFHMNSNQIK 60
 DB 1 MNRGVPFRHLLVQLALPAAATQGNKVLGKGGDTVELTCTASOKKSIQFHMNSNQIK 60
 QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
 DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120
 QY 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180
 DB 121 LVFGLTANSPTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180
 QY 181 TWCTCTVLQNGKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
 DB 181 TWCTCTVLQNGKVEFKIDIVLAFOKASSIVYKKEGEQVEFSPPLAFTVEKLTGSGELMW 240
 QY 241 QAEBAASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPULHLPQALPOYAGSGNLTLA 300
 DB 241 QAEBAASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPULHLPQALPOYAGSGNLTLA 300
 QY 301 LEAKTGKLGHOENVLVMMARATQLOKNLICEVWGPTSPKMLSLKENKAIVSKREKPVW 360
 DB 301 LEAKTGKLGHOENVLVMMARATQLOKNLICEVWGPTSPKMLSLKENKAIVSKREKPVW 360
 QY 211 -----EC-----PCPAPP-----VAGSVF 226
 DB 211 -----EC-----PCPAPP-----VAGSVF 226
 QY 421 LPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPREQVYTLR 480
 DB 421 LPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPREQVYTLR 480
 QY 227 LPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPREQVYTLR 286
 DB 227 LPPPKDITLMSRTPEVTCVVDVSHEDPEVKFMYVDGVEVHNAKTKPREQVYTLR 286
 QY 481 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLRPSRDELTKN 540
 DB 481 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLRPSRDELTKN 540
 QY 287 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLRPSRDELTKN 346
 DB 287 VSVLTVLHODWLNKKEYCKVSNKALPAPIEKTISKAKGQPREQVYTLRPSRDELTKN 346
 QY 541 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTVDSRMOQGN 600
 DB 541 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTVDSRMOQGN 600
 QY 347 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTVDSRMOQGN 406
 DB 347 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPVLDSDGFFLYSKLTVDSRMOQGN 406

QY 601 VFSCSVHHEALHNHYTOKSLSPG 625
DB 407 VFSCSVHHEALHNHYTOKSLSPG 431

RESULT 4

US-09-766-995-2
Sequence 2, Application US/09766995
Patent No. US20020052481A1
GENERAL INFORMATION:
APPLICANT: Graham P. Allaway et al.
TITLE OF INVENTION: NON-PEPTIDYL MOIETY-CONJUGATED CD4-GAMMA2 AND CD4-1G62 IMMUNOCONJ
FILE REFERENCE: 2048/41215-CB/JPM/SHS
CURRENT APPLICATION NUMBER: US/09/766,995
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 432
TYPE: PRT
ORGANISM: homo sapiens
US-09-766-995-2

Query Match 60.3%; Score 2059; DB 9; Length 432;
Best Local Similarity 65.6%; Pred. No. 1,4e-137;

Matches 410; Conservative 12; Mismatches 9; Indels 194; Gaps 3;

QY 1 MNRGVFPRHLLVLTQALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHKNSNOIK 60
DB 1 MNRGVFPRHLLVLTQALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHKNSNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKIKIEDSDTYICEVEDQKEEVOJ 120
DB 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKIKIEDSDTYICEVEDQKEEVOJ 120
QY 121 LVFGLTANSPTHLQOGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
DB 121 LVFGLTANSPTHLQOGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
QY 181 TWICTVLOQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPFLAFTVEKLTGSGELMW 240
DB 181 TWICTVLOQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPFLAFTVEKLTGSGELMW 240
QY 241 QAERASSKSWITFDLKNKEVSVKRVTDPKLOMGKKLPLHLTPQALPOYAGSGMLTLA 300
DB 241 QAERASSKSWITFDLKNKEVSVKRVTDPKLOMGKKLPLHLTPQALPOYAGSGMLTLA 300
QY 301 LEAKTGKHQEVNLVVMRATQLOKNLTCCEVWGPISPKMLSLKLENKEAKVSKREKPVW 360
DB 301 LEAKTGKHQEVNLVVMRATQLOKNLTCCEVWGPISPKMLSLKLENKEAKVSKREKPVW 360
QY 421 LPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNMTYDGEVHNATKPREEOYNSTR 480
DB 421 LPPPKKDTLMISRTPEVTCVVDVSHEDPEVKFNMTYDGEVHNATKPREEOYNSTR 480
QY 227 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMTYDGEVHNATKPREEOYNSTR 286
DB 227 LFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMTYDGEVHNATKPREEOYNSTR 286
QY 481 VVSVLTVLHODMNGKEKCKVSKKALPAPIEKTIKSAKGPPEPOVYTLPPSRDELTKN 540
DB 481 VVSVLTVLHODMNGKEKCKVSKKALPAPIEKTIKSAKGPPEPOVYTLPPSRDELTKN 540
QY 287 VVSVLTVLHODMNGKEKCKVSKKALPAPIEKTIKSAKGPPEPOVYTLPPSRDELTKN 346
DB 287 VVSVLTVLHODMNGKEKCKVSKKALPAPIEKTIKSAKGPPEPOVYTLPPSRDELTKN 346
QY 541 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 600
DB 541 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 600
QY 347 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 406
DB 347 QVSLTCLVKGFPYPSDIAVEMESNGOPENNYKTPPVLDSDGSPFLYSKLTVDKSRMOQGN 406
QY 601 VFSCSVHHEALHNHYTOKSLSPG 625
DB 407 VFSCSVHHEALHNHYTOKSLSPG 431

RESULT 5

US-09-939-537-6
Sequence 6, Application US/09939537
Patent No. US20030138410A1
GENERAL INFORMATION:
APPLICANT: Seed, Brian

Banapur, Babak
Romeo, Charles
Kolatus, Waldeemar

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSER: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,537
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284,391
FILING DATE: 02-Aug-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045

TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-939-537-6

Query Match 59.7%; Score 2036.5; DB 10; Length 532;
Best Local Similarity 98.3%; Pred. No. 7.3e-136;

Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVFPRHLLVLTQALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHKNSNOIK 60
DB 1 MNRGVFPRHLLVLTQALPAATQGNKVVLGKGGDTVELTCTASOKKSIOFHKNSNOIK 60
QY 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKIKIEDSDTYICEVEDQKEEVOJ 120
DB 61 ILGNQGSFLTKGPKSKLNDRAISRSLMDQGNFPLIINKIKIEDSDTYICEVEDQKEEVOJ 120
QY 121 LVFGLTANSPTHLQOGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
DB 121 LVFGLTANSPTHLQOGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELODSG 180
QY 181 TWICTVLOQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPFLAFTVEKLTGSGELMW 240
DB 181 TWICTVLOQKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPFLAFTVEKLTGSGELMW 240

Db 181 TWCTVLOKQKVEKIDIVLAFQKASSIVYKKEGQVEFSPLAFVTEKLTSGELMW 240
Qy 241 QAEARASSKSWITPDLKNKEVSVKRVYODPDLQWCKLPLHLTLPQALPOYAGSGNLTLA 300
Db 241 QAEARASSKSWITPDLKNKEVSVKRVYODPDLQWCKLPLHLTLPQALPOYAGSGNLTLA 300
Qy 301 LEATGKLGHOEVNLVVRATOLQKNTCEVWGPTSPKMLSLKLENKAKYSKREKPYWV 360
Db 301 LEATGKLGHOEVNLVVRATOLQKNTCEVWGPTSPKMLSLKLENKAKYSKREKPYWV 360
Qy 361 LNPEAGMWQCLISDSGQVLLSNIKVLPTWSTPYPCAPBPCKSC 404
Db 361 LNPEAGMWQCLISDSGQVLLSNIKVLPTWSTPYVHA---DPKLC 401

RESULT 6
US-09-243-008-6
; Sequence 6, Application US/09243008
; Publication No. US2004000534A1
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; Receptor Chimeras
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/243,008
; FILING DATE: 02-Feb-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,176
; FILING DATE: SEPTEMBER 11, 1995
; APPLICATION NUMBER: 08/203,866
; FILING DATE: February 28, 1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: March 6, 1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: March 7, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Karen F. Lech, Ph.D.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/270001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-243-008-6

Query Match 59.7%; Score 2036.5; DB 11; Length 532;
Best Local Similarity 98.3%; Pred. No. 7.3e-116;
Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Qy 1 MNRGVPFRHLLVLQALLPATQGNKVLKGGDVELTCTASQKSIQPHMKNNOIK 60
Db 1 MNRGVPFRHLLVLQALLPATQGNKVLKGGDVELTCTASQKSIQPHMKNNOIK 60

Qy 61 ILGNQGSFLTRKPSKLNDRADSRRLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVL 120
Db 61 ILGNQGSFLTRKPSKLNDRADSRRLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVL 120
Qy 121 LVFGLTANSDTHLLQGGSLTFLTESPPGSSPSVQCRSPKGNIOGKTLISVQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGGSLTFLTESPPGSSPSVQCRSPKGNIOGKTLISVQLELDQSG 180
Qy 181 TWCTVLOKQKVEKIDIVLAFQKASSIVYKKEGQVEFSPLAFVTEKLTSGELMW 240
Db 181 TWCTVLOKQKVEKIDIVLAFQKASSIVYKKEGQVEFSPLAFVTEKLTSGELMW 240
Qy 241 QAEARASSKSWITPDLKNKEVSVKRVYODPDLQWCKLPLHLTLPQALPOYAGSGNLTLA 300
Db 241 QAEARASSKSWITPDLKNKEVSVKRVYODPDLQWCKLPLHLTLPQALPOYAGSGNLTLA 300
Qy 301 LEATGKLGHOEVNLVVRATOLQKNTCEVWGPTSPKMLSLKLENKAKYSKREKPYWV 360
Db 301 LEATGKLGHOEVNLVVRATOLQKNTCEVWGPTSPKMLSLKLENKAKYSKREKPYWV 360
Qy 361 LNPEAGMWQCLISDSGQVLLSNIKVLPTWSTPYPCAPBPCKSC 404
Db 361 LNPEAGMWQCLISDSGQVLLSNIKVLPTWSTPYVHA---DPKLC 401

RESULT 7
US-09-939-537-4
; Sequence 4, Application US/09939537
; Publication No. US20030138410A1
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; Banapour, Babak
; Romeo, Charles
; Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; CELLS BY CHIMERIC CD4 RECEPTOR - BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,537
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

LENGTH: 575 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-939-537-4

Query Match 59.7%; Score 2036.5; DB 10; Length 575;
 Best Local Similarity 98.3%; Pred. No. 8.1e-136;
 Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
 DB 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
 QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNPFLIIKNLIKEDSDTYICEVEDQKEEYVL 120
 DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNPFLIIKNLIKEDSDTYICEVEDQKEEYVL 120
 QY 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
 DB 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
 QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFSFPLAFVTEKLTGSGELMW 240
 DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFSFPLAFVTEKLTGSGELMW 240
 QY 241 QARRASSSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
 DB 241 QARRASSSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
 QY 301 LEAKTGKLEHGVNLVVRATQLOKNTLCEVWGPTSPKMLSLKLENKAVSREKRPVWY 360
 DB 301 LEAKTGKLEHGVNLVVRATQLOKNTLCEVWGPTSPKMLSLKLENKAVSREKRPVWY 360
 QY 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVCAPAPKSC 404
 DB 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVCAPAPKSC 404

RESULT 8
 US-09-243-008-4
 Sequence 4, Application US/09243008
 Publication No. US20040005334A1
 GENERAL INFORMATION:
 APPLICANT: Seed, Brian et al.
 TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor Chimeras
 NUMBER OF SEQUENCES: 40
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 502 or SSSX
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 SOFTWARE: Wordperfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/243.008
 FILING DATE: 02-Feb-1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/394.176
 FILING DATE: SEPTEMBER 11, 1995
 APPLICATION NUMBER: 08/203,866
 FILING DATE: February 28, 1994
 APPLICATION NUMBER: 07/847,566
 FILING DATE: March 6, 1992
 APPLICATION NUMBER: 07/665,961

FILING DATE: March 7, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Karen F. Lech, Ph.D
 REGISTRATION NUMBER: 35,238
 REFERENCE/DOCKET NUMBER: 00786/270001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070
 TELEFAX: (617) 542-8906
 TELE: 200154
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 575 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-243-008-4

Query Match 59.7%; Score 2036.5; DB 11; Length 575;
 Best Local Similarity 98.3%; Pred. No. 8.1e-136;
 Matches 397; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
 DB 1 MNRGVPFRHLVLVQLALPAATQGNKVVLGKKGDVVELTCTASQKKSIOFHMKNSNQIK 60
 QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNPFLIIKNLIKEDSDTYICEVEDQKEEYVL 120
 DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLMDQGNPFLIIKNLIKEDSDTYICEVEDQKEEYVL 120
 QY 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
 DB 121 LVFGLTANSDTHLLQGSQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLSSVSOLELQDSG 180
 QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFSFPLAFVTEKLTGSGELMW 240
 DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKASSIVYKKEGEVFSFPLAFVTEKLTGSGELMW 240
 QY 241 QARRASSSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
 DB 241 QARRASSSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPQALPOYAGSGNLTLLA 300
 QY 301 LEAKTGKLEHGVNLVVRATQLOKNTLCEVWGPTSPKMLSLKLENKAVSREKRPVWY 360
 DB 301 LEAKTGKLEHGVNLVVRATQLOKNTLCEVWGPTSPKMLSLKLENKAVSREKRPVWY 360
 QY 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVCAPAPKSC 404
 DB 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPVCAPAPKSC 404

RESULT 9
 US-09-939-537-5
 Sequence 5, Application US/0939537
 Publication No. US20030138410A1
 GENERAL INFORMATION:
 APPLICANT: Seed, Brian
 Banapour, Babak
 Romeo, Charles
 Kolanus, Waldemar
 TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
 NUMBER OF SEQUENCES: 53
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Clark & Elbing LLP
 STREET: 176 Federal Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/939,537
FILING DATE: 24-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-939-537-5

Query Match 59.5%; Score 2032.5; DB 10; Length 462;
Best Local Similarity 98.0%; Pred. No. 1.2e-135;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
QY 1 MNRGVPFRLHLLVQLALLPAATQGNKRVLGKGGDTVELTCTASQKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRLHLLVQLALLPAATQGNKRVLGKGGDTVELTCTASQKSIQFHMKNNSQIK 60
QY 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKI ESDTYICEVEDQKEEYQL 120
DB 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKI ESDTYICEVEDQKEEYQL 120
QY 121 LVFGITANSPTHLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
DB 121 LVFGITANSPTHLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
QY 181 TWTCVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
DB 181 TWTCVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPLQALPOYAGSGNLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPLQALPOYAGSGNLTLA 300
QY 301 LEATGKLHGEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
DB 301 LEATGKLHGEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
QY 361 LNPEAGMOCILSDSGOVLLESNIKVLPTWSTPVCAPPEKSC 404
DB 361 LNPEAGMOCILSDSGOVLLESNIKVLPTWSTPVCAPPEKSC 404

RESULT 10
US-09-243-008-5
Sequence 5, Application US/09243008
Publication No. US2004000534A1
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by

Receptor Chimeras
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/243,008
FILING DATE: 02-Feb-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/394,176
FILING DATE: SEPTEMBER 11, 1995
APPLICATION NUMBER: 08/203,866
FILING DATE: February 28, 1994
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Karen F. Lech, Ph.D.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/270001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-243-008-5
Query Match 59.5%; Score 2032.5; DB 11; Length 462;
Best Local Similarity 98.0%; Pred. No. 1.2e-135;
Matches 396; Conservative 2; Mismatches 3; Indels 3; Gaps 1;
QY 1 MNRGVPFRLHLLVQLALLPAATQGNKRVLGKGGDTVELTCTASQKSIQFHMKNNSQIK 60
DB 1 MNRGVPFRLHLLVQLALLPAATQGNKRVLGKGGDTVELTCTASQKSIQFHMKNNSQIK 60
QY 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKI ESDTYICEVEDQKEEYQL 120
DB 61 ILGNQSFLLTKGPKSLNDRADSRSLMDQGNFPLIINKLKI ESDTYICEVEDQKEEYQL 120
QY 121 LVFGITANSPTHLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
DB 121 LVFGITANSPTHLQGGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDDSG 180
QY 181 TWTCVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
DB 181 TWTCVLOKQKVEFKIDIVLAFQKASSIYKKKEGQVEFSPFLATFVEKLTGSGELMW 240
QY 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPLQALPOYAGSGNLTLA 300
DB 241 QAERASSSKSWITFDLKNKEVSVKRVTDPKLQMGKPLHLTLPLQALPOYAGSGNLTLA 300
QY 301 LEATGKLHGEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
DB 301 LEATGKLHGEVNLVVMRATOLQKNLTCEVWGPTSPKMLSLKLENKEAVSKREKPYWV 360
QY 361 LNPEAGMOCILSDSGOVLLESNIKVLPTWSTPVCAPPEKSC 404
DB 361 LNPEAGMOCILSDSGOVLLESNIKVLPTWSTPVCAPPEKSC 404


```

; CURRENT APPLICATION NUMBER: US/10/188,444
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-444-39

Query Match      59.0%; Score 2015; DB 14; Length 458;
Best Local Similarity 99.5%; Pred. No. 2e-134;
Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALALPAATQGNKVLGKKGDVVELTCTASQKKSIQFHMKNNOIK 60
DB 1 MNRGVPRHLLVLTQALALPAATQGNKVLGKKGDVVELTCTASQKKSIQFHMKNNOIK 60

QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120

QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWCTVLOQOKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQOKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240

QY 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300
DB 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300

QY 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300
DB 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300

QY 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAKVSKREKPVWY 360
DB 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAKVSKREKPVWY 360

QY 361 LNPEAGMWQCLLSDSGQVLLSENIKVLPTWSTPV 394
DB 361 LNPEAGMWQCLLSDSGQVLLSENIKVLPTWSTPV 394

RESULT 16
US-10-207-655-170
; Sequence 170, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069 401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-170

Query Match      59.0%; Score 2015; DB 14; Length 458;
Best Local Similarity 99.5%; Pred. No. 2e-134;
Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALALPAATQGNKVLGKKGDVVELTCTASQKKSIQFHMKNNOIK 60
DB 1 MNRGVPRHLLVLTQALALPAATQGNKVLGKKGDVVELTCTASQKKSIQFHMKNNOIK 60
```

```

QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120

QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWCTVLOQOKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQOKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240

QY 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300
DB 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300

QY 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAKVSKREKPVWY 360
DB 301 LEAKTGKLEHVEVNLVYMRATQLOKNTLCEVWGPTSPKMLSLKLENKEAKVSKREKPVWY 360

QY 361 LNPEAGMWQCLLSDSGQVLLSENIKVLPTWSTPV 394
DB 361 LNPEAGMWQCLLSDSGQVLLSENIKVLPTWSTPV 394

RESULT 17
US-10-151-274-3
; Sequence 3, Application US/10151274
; Publication No. US20030064071A1
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geijlendeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY
; TITLE OF INVENTION: INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/10/151,274
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: US/09/517,605
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-151-274-3

Query Match      58.8%; Score 2007; DB 12; Length 458;
Best Local Similarity 99.2%; Pred. No. 7.4e-134;
Matches 391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALALPAATQGNKVLGKKGDVVELTCTASQKKSIQFHMKNNOIK 60
DB 1 MNRGVPRHLLVLTQALALPAATQGNKVLGKKGDVVELTCTASQKKSIQFHMKNNOIK 60

QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120
DB 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEVQL 120

QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWCTVLOQOKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
DB 181 TWCTVLOQOKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240

QY 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300
DB 241 QAEARASSKSWITFDLKNKEVSVKRYTQDPKLOMGKKLPLHLTLPALPOYAGSGNLTLLA 300
```

Db 241 QAERASSSKSWITFDLKNKEVSVKRVYDTPKLOMGKLLPLHLTLPOALPOYAGSGNLTIA 300
 QY 301 LEAKTGKHOEVNLVVMKATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYV 360
 Db 301 LEAKTGKHOEVNLVVMKATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKAVV 360
 QY 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPV 394
 Db 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPV 394

RESULT 18

US-09-891-119A-2

Sequence 2, Application US/09891119A

Publication No. US20040013683A1

GENERAL INFORMATION:

APPLICANT: Maddon, Paul J.

TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4

FILE REFERENCE: 24577-CY-B

CURRENT APPLICATION NUMBER: US/09/891,119A

CURRENT FILING DATE: 2001-06-25

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 2

LENGTH: 397

TYPE: PRT

ORGANISM: Human

US-09-891-119A-2

Query Match

58.6%; Score 2001; DB 11; Length 397;

Best Local Similarity 98.7%; Pred. No. 1.6e-133;

Matches 389; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKGDVLELTCTASQKKSIOFHMKNNOIK 60
 Db 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKGDVLELTCTASQKKSIOFHMKNNOIK 60
 QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120
 Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120
 QY 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 Db 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 QY 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 Db 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 QY 181 TWCTCVLQONKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 Db 181 TWCTCVLQONKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 QY 241 QAERASSSKSWITFDLKNKEVSVKRVYDTPKLOMGKLLPLHLTLPOALPOYAGSGNLTIA 300
 Db 241 QAERASSSKSWITFDLKNKEVSVKRVYDTPKLOMGKLLPLHLTLPOALPOYAGSGNLTIA 300
 QY 301 LEAKTGKHOEVNLVVMKATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYV 360
 Db 301 LEAKTGKHOEVNLVVMKATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKAVV 360
 QY 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPV 394
 Db 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPV 394

RESULT 19

US-08-681-219-27

Sequence 27, Application US/08681219

Publication No. US20020058607A1

GENERAL INFORMATION:

APPLICANT: Takeaki Sato and Junn Yanagisawa

TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN

TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLCF

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/681,219
 FILING DATE: 22-Jul-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JKM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 458 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-681-219-27

Query Match

58.5%; Score 1997; DB 8; Length 458;

Best Local Similarity 98.7%; Pred. No. 3.8e-133;

Matches 389; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKGDVLELTCTASQKKSIOFHMKNNOIK 60
 Db 1 MNRGVPFRHLLVYLQALLPAATQGNKVVLGKGDVLELTCTASQKKSIOFHMKNNOIK 60
 QY 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120
 Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEVQL 120
 QY 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 Db 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 QY 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 Db 121 LVFGLTANSPTHLLOQGSLLTLTLESPPGSSPVQCRSPRGKNIQSGKTLVSQLELDDSG 180
 QY 181 TWCTCVLQONKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 Db 181 TWCTCVLQONKQKVEFKIDIVLAFQKASSIVYKKEGEQVEFSPLAFTVEKLTGSGELMW 240
 QY 241 QAERASSSKSWITFDLKNKEVSVKRVYDTPKLOMGKLLPLHLTLPOALPOYAGSGNLTIA 300
 Db 241 QAERASSSKSWITFDLKNKEVSVKRVYDTPKLOMGKLLPLHLTLPOALPOYAGSGNLTIA 300
 QY 301 LEAKTGKHOEVNLVVMKATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKPYV 360
 Db 301 LEAKTGKHOEVNLVVMKATOLQKNLTCEVWGPTSPKMLSLKLENKAKVSKREKAVV 360
 QY 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPV 394
 Db 361 LNPEAGMWQCLLSDSGQVLLSNIKVLPTWSTPV 394

RESULT 20

US-09-230-111C-25

Sequence 25, Application US/09230111C

Publication No. US20030203414A1

GENERAL INFORMATION:

APPLICANT: Sato, Taka-Aki

TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN

TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLCF (PDZ/DHR)

CORRESPONDENCE ADDRESS:

;; TITLE OF INVENTION: DOMAIN AND USES THEREOF
;; FILE REFERENCE: 48963-A-PCF-US
;; CURRENT APPLICATION NUMBER: US/09/230,111C
;; CURRENT FILING DATE: 1999-05-17
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 25
;; LENGTH: 458
;; TYPE: PRP
;; ORGANISM: human
US-09-230-111C-25

Query Match 58.5%; Score 1997; DB 11; Length 458;
Best Local Similarity 98.7%; Pred. No. 3.8e-133;
Matches 389; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLLVQLALPPAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPRHLLLVQLALPPAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
QY 61 ILGNQGSFLTKGSKLNDRAISRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSKLNDRAISRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPVQCRSPRGKNIQGGKTLISVQLBLQDSG 180
DB 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPVQCRSPRGKNIQGGKTLISVQLBLQDSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QARRASSKSWITFDLKNKEVSVKRVTDPKLQMKKPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QARRASSKSWITFDLKNKEVSVKRVTDPKLQMKKPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLGHOENVLVVWRATQLOKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPVVY 360
DB 301 LEAKTGKLGHOENVLVVWRATQLOKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPVVY 360
QY 361 LNPEAGMOCCLSDSGOVLLESNINIKVLPWTSTPV 394
DB 361 LNPEAGMOCCLSDSGOVLLESNINIKVLPWTSTPV 394

RESULT 21
US-10-092-138-25
;; Sequence 25, Application US/10092138
;; Publication No. US20030170723A1
;; GENERAL INFORMATION:
;; APPLICANT: Sato, Taka-Aki
;; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
;; FILE REFERENCE: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
;; CURRENT APPLICATION NUMBER: US/10/092,138
;; CURRENT FILING DATE: 2002-09-06
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 25
;; LENGTH: 458
;; TYPE: PRP
;; ORGANISM: human
US-10-092-138-25

Query Match 58.5%; Score 1997; DB 14; Length 458;
Best Local Similarity 98.7%; Pred. No. 3.8e-133;
Matches 389; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLLVQLALPPAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPRHLLLVQLALPPAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
QY 61 ILGNQGSFLTKGSKLNDRAISRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120

DB 61 ILGNQGSFLTKGSKLNDRAISRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPVQCRSPRGKNIQGGKTLISVQLBLQDSG 180
DB 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPVQCRSPRGKNIQGGKTLISVQLBLQDSG 180
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKASSIVYKKEGQVEFSPPLAFTVEKLTGSGELMW 240
QY 241 QARRASSKSWITFDLKNKEVSVKRVTDPKLQMKKPLHLTLPOALPOYAGSGNLTIA 300
DB 241 QARRASSKSWITFDLKNKEVSVKRVTDPKLQMKKPLHLTLPOALPOYAGSGNLTIA 300
QY 301 LEAKTGKLGHOENVLVVWRATQLOKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPVVY 360
DB 301 LEAKTGKLGHOENVLVVWRATQLOKNLTCEVWGPTSPKMLSLKENKEAKVSKREKPVVY 360
QY 361 LNPEAGMOCCLSDSGOVLLESNINIKVLPWTSTPV 394
DB 361 LNPEAGMOCCLSDSGOVLLESNINIKVLPWTSTPV 394

RESULT 22
US-10-097-044A-4
;; Sequence 4, Application US/10097044A
;; Publication No. US20030143220A1
;; GENERAL INFORMATION:
;; APPLICANT: Capon, Daniel J.
;; Gregory, Timothy J.
;; TITLE OF INVENTION: Adhesion Variants
;; NUMBER OF SEQUENCES: 25
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Genentech, Inc.
;; STREET: 460 Point San Bruno Blvd
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patin (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/097,044A
;; FILING DATE: 28-May-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/457,918
;; FILING DATE: 1-JUN-1995
;; APPLICATION NUMBER: 08/236311
;; FILING DATE: 02-MAY-1994
;; APPLICATION NUMBER: 07/936190
;; FILING DATE: 26-AUG-1992
;; APPLICATION NUMBER: 07/842777
;; FILING DATE: 18-FEB-1992
;; APPLICATION NUMBER: 07/250785
;; FILING DATE: 28-SEP-1988
;; APPLICATION NUMBER: 07/104329
;; FILING DATE: 02-OCT-1987
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kudinec, Jeffrey S.
;; REGISTRATION NUMBER: 36,575
;; REFERENCE/DOCKET NUMBER: P0444P1C3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-8228
;; TELEFAX: 415/953-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 434 amino acids

TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-097-044A-4

Query Match 55.8%; Score 1904; DB 14; Length 434;
Best Local Similarity 99.7%; Pred. No. 1.4e-126;
Matches 369; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 QGNKVLGKKDPTVELTCTASQKSIQFHMKNNSNOIKILNQGSLTGTGPKLNDRAISR 83
DB 56 QGNKVLGKKDPTVELTCTASQKSIQFHMKNNSNOIKILNQGSLTGTGPKLNDRAISR 115
QY 84 RSLMDQGNFPLIINKLTIEDSDTYICEVEDQKEEYQVLLVFGLTANSPTHLLOQGSLLTTL 143
DB 116 RSLMDQGNFPLIINKLTIEDSDTYICEVEDQKEEYQVLLVFGLTANSPTHLLOQGSLLTTL 175
QY 144 ESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLQNKKEVFKIDIVVLA 203
DB 176 ESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLQNKKEVFKIDIVVLA 235
QY 204 FOKASSIYKKEGEQVEFSFPLAFVETLTSGEELMWQABRASSSKSWITFDLKNKEYSV 263
DB 236 FOKASSIYKKEGEQVEFSFPLAFVETLTSGEELMWQABRASSSKSWITFDLKNKEYSV 295
QY 264 KRVTDPKLQMGKKLPLHLTLPOALPOYAGSGNLTLEAKTGKLDHENVLVWRATQLO 323
DB 296 KRVTDPKLQMGKKLPLHLTLPOALPOYAGSGNLTLEAKTGKLDHENVLVWRATQLO 355
QY 324 KNLTCFVWGPTSPKMLSLKLENKAKVSKREKPVWVLPENAGMWQCLLSDSGVLLSESN 383
DB 356 KNLTCFVWGPTSPKMLSLKLENKAKVSKREKPVWVLPENAGMWQCLLSDSGVLLSESN 415
QY 384 IKVLPWSTP 393
DB 416 IKVLPWSTP 425

RESULT 23
US-09-759-841-6
Sequence 6, Application US/09759841
Patent No. US20010039026A1
GENERAL INFORMATION:
APPLICANT: Rickett, Graham A
APPLICANT: Dobbs, Susan
APPLICANT: Petrov, Manoussos
TITLE OF INVENTION: Assay Method
FILE REFERENCE: PCI0348APME
CURRENT APPLICATION NUMBER: US/09/759,841
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: GB 0000661.9
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: GB 0000663.5
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: GB 0000659.3
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
US-09-759-841-6

Query Match 55.4%; Score 1891; DB 9; Length 370;
Best Local Similarity 99.7%; Pred. No. 9.3e-126;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 27 KVLGKKDDTVELTCTASQKSIQFHMKNNSNOIKILNQGSLTGTGPKLNDRAISRSL 86
DB 2 KVLGKKDDTVELTCTASQKSIQFHMKNNSNOIKILNQGSLTGTGPKLNDRAISRSL 61
QY 87 WDOGNFPLIINKLTIEDSDTYICEVEDQKEEYQVLLVFGLTANSPTHLLOQGSLLTLESP 146

DB 62 WDOGNFPLIINKLTIEDSDTYICEVEDQKEEYQVLLVFGLTANSPTHLLOQGSLLTLESP 121
QY 147 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLQNKKEVFKIDIVVLAFOK 206
DB 122 PGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWCTVLQNKKEVFKIDIVVLAFOK 181
QY 207 ASSIYKKEGEQVEFSFPLAFVETLTSGEELMWQABRASSSKSWITFDLKNKEYSYKR 241
DB 182 ASSIYKKEGEQVEFSFPLAFVETLTSGEELMWQABRASSSKSWITFDLKNKEYSYKR 241
QY 267 TQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLEAKTGKLDHENVLVWRATQLOKNT 326
DB 242 TQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTLEAKTGKLDHENVLVWRATQLOKNT 301
QY 327 TCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPENAGMWQCLLSDSGVLLSESNKY 386
DB 302 TCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPENAGMWQCLLSDSGVLLSESNKY 361
QY 387 LPTWSTPV 394
DB 362 LPTWSTPV 369

RESULT 24
US-10-024-329-32
Sequence 32, Application US/10024329
Publication No. US20030157063A1
GENERAL INFORMATION:
APPLICANT: SANHADI, Kamel
APPLICANT: TOURAINE, Jean-Louis
APPLICANT: LEROY, Pierre
APPLICANT: MERTALI, Majid
TITLE OF INVENTION: Gene therapy using anti-gp41 antibody and cd4 immunoadhesin
FILE REFERENCE: 109993
CURRENT APPLICATION NUMBER: US/10/024,329
CURRENT FILING DATE: 2001-12-21
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 32
LENGTH: 448
TYPE: PRT
ORGANISM: human scd4
US-10-024-329-32

Query Match 55.0%; Score 1877.5; DB 14; Length 448;
Best Local Similarity 97.0%; Pred. No. 1.1e-124;
Matches 382; Conservative 0; Mismatches 3; Indels 9; Gaps 8;

QY 1 MNRGVPRHLLVLVQLALLPAATQGNKVLGKKDPTVELTCTASQKSIQFHMKNNSNOIK 60
DB 1 MNRGVPRHLLVLVQLALLPAATQGNKVLGKKDPTVELTCTASQKSIQFHMKNNSNOIK 59
QY 61 ILNQGSLTGTGPKLNDRAISRSLMDQGNFPLIINKLTIEDSDTYICEVEDQKEEYQV 120
DB 60 ILNQGSLTGTGPKLNDRAISRSLMDQGNFPLIINKLTIEDSDTYICEVEDQKEEYQV 117
QY 121 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 118 LVFGLTANSPTHLLOQGSLLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 174
QY 181 TWCTCTVLQNKKEVFKIDIVVLAFOKASSIYKKEGEQVEFSFPLAFVETLTSGEELMW 240
DB 175 TWCTCTVLQNKKEVFKIDIVVLAFOKASSIYKKEGEQVEFSFPLAFVETLTSGEELMW 231
QY 241 QABRASSSKSWITFDLKNKEYSYKRVTQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTTLA 300
DB 232 QABRASSSKSWITFDLKNKEYSYKRVTQDPKLOMGKKLPLHLTLPOALPOYAGSGNLTTLA 291
QY 301 LEAKTGKLDHENVLVWRATQLOKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPVWV 360
DB 292 LEAKTGKLDHENVLVWRATQLOKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPVWV 351

QY 361 LNPEAGMOCCLSDSGOVLLESNIKVLPTMTSPV 394
DB 352 LNPEAGMOCCLSDSGOVLLESNIKVLPTMTSPV 385

RESULT 25

US-09-939-537-33

Sequence 33, Application US/09939537

Publication No. US20030138410A1

GENERAL INFORMATION:

APPLICANT: Seed, Brian

Banapur, Babak

Romeo, Charles

Kolanus, Waldemar

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED

CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESS: Clark & Elbing LLP

STREET: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/939,537

FILING DATE: 24-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/284,391

FILING DATE: 02-AUG-1994

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 254 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 33:

US-09-939-537-33

Query Match 39.2%; Score 1338.5; DB 10; Length 254;

Best Local Similarity 98.0%; Pred. No. 8.5e-87; Indels 5; Gaps 1;

Matches 249; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 400 EPKSCDKTHTC-----PELLGSPVFLFPPKPKDTLMISRTPEVTCVAVVSHEDPEVKF 454

DB 1 EPKSCDKTHTCPRPAPELLGSPVFLFPPKPKDTLMISRTPEVTCVAVVSHEDPEVKF 60

QY 455 NMVYDGVVNAKTKPREBOYNSTYRVVSVLTIVLHQMNLGKCYKCVSKALPAPIEKT 514

DB 61 NMVYDGVVNAKTKPREBOYNSTYRVVSVLTIVLHQMNLGKCYKCVSKALPAPIEKT 120

QY 515 ISKAKGQRPREQVYTLPPSRDELTKNQVSLTCLVKGYFPPSDINAVESNGQPENNYKTP 574

DB 121 ISKAKGQRPREQVYTLPPSRDELTKNQVSLTCLVKGYFPPSDINAVESNGQPENNYKTP 180

QY 575 PVLDSGSPFLYSKLTVDKSRWQGNVFSVYMEALNNHYTKSLSPGLQDETCAE 634

DB 181 PVLDSGSPFLYSKLTVDKSRWQGNVFSVYMEALNNHYTKSLSPGLQDETCAE 240

QY 635 AODGELDGLWTTDP 648

DB 241 AODGELDGLWTTDP 254

RESULT 26

US-10-363-427-18

Sequence 18, Application US/10363427

Publication No. US2003019538A1

GENERAL INFORMATION:

APPLICANT: Medexgen Inc.

APPLICANT: CHUNG, Yong Hoon

APPLICANT: HAN, Ji Woong

APPLICANT: LEE, Hye Ja

APPLICANT: CHOI, Eun Yong

APPLICANT: KIM, Jin Mi

APPLICANT: YIM, Soo Bin

TITLE OF INVENTION: Concatameric Immunoadhesion

FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/363,427

CURRENT FILING DATE: 2003-02-28

NUMBER OF SEQ ID NOS: 52

SOFTWARE: Kopatentin 1.71

SEQ ID NO 18

LENGTH: 617

TYPE: PRT

ORGANISM: Homo sapiens

US-10-363-427-18

Query Match 38.5%; Score 1313.5; DB 14; Length 617;

Best Local Similarity 49.0%; Pred. No. 1.7e-84; Indels 141; Gaps 24;

Matches 317; Conservative 52; Mismatches 137; Indels 141; Gaps 24;

QY 35 DIVERLCTSAQKKSIOFHKNKSNQIKILGNQGSFLTQKPSKLNDRASRLMDQGNFPL 94

DB 55 DDKMEKTSKKKIAQFRKEKE-----TFKEDTYKLFK-----NGTL 92

QY 95 IIKNLKIEDSDTYICEVEQK-EVQVLVFGLTANSDTLLDQO-----SLTLT 142

DB 93 KIKHLKTDQDIIYKVSIVDTKGNVLEKIFDLK-----IQERSKPKISWTCINTTLT 145

QY 143 LSSPSSPSVOCSPRGKNIQGGKTLVSQLELQDSGTWCTCVLQNKVEFKIDIVL 202

DB 146 CEVNGTDBELNL-----YDGHKLKLSQVIT--THKNTLSL-----SAFK-----CT 187

QY 203 AFQKASIVYKKEGROVESFPLAFTVEKLTGSGELM----- 239

DB 188 AGNKVS-----KESSVEPVSCP-----KEITVALETWGALGODINDIPSPQWSDIDDI 237

QY 240 -WOERASSSKSWITFD-----LAKKEVSKRYTQPKLQMKKLPALHL 284

DB 238 KW-EKTSIDKIKIAQFRKEKEFTKEDTYKLFKNGTLKIKHLKTDQ--ODIYKVSIVDT- 292

QY 285 PQLPQVAGSNVLTALAEKTKLHGVNLVVMRATOLQKNTLCEVWGSPPLMLSLK 344

DB 293 -----KKNVLEKIFDL-----KIQERSKPKISWTCINTTLCEVNGTDBE--LNLVQ 340

QY 345 ENKEAKVSKRE-KPVVNLNPEAGMOCCLSDSGOVLLESNIKVLPTWSTPVPAPPEK 403

DB 341 DGHKLKLSQVITTHKNTLSLA-KFKC--TAKNKVSKSSV-----PVSCPA-EPKS 389

QY 404 CDKTHTC-----PELLGSPVFLFPPKPKDTLMISRTPEVTCVAVVSHEDPEVK 458

DB 390 CDKTHTCPPCPAPPELLGSPVFLFPPKPKDTLMISRTPEVTCVAVVSHEDPEVK 449

QY 459 DGVVNAKTKPREBOYNSTYRVVSVLTIVLHQMNLGKCYKCVSKALPAPIEKTISKA 518


```

1  CURRENT APPLICATION NUMBER: US/09/933,868
2  CURRENT FILING DATE: 2002-04-11
3  PRIOR APPLICATION NUMBER: PCT/US99/22045
4  PRIOR FILING DATE: 1999-09-22
5  NUMBER OF SEQ ID NOS: 52
6  SOFTWARE: PatentIn version 3.0
7  SEQ ID NO 8
8  LENGTH: 592
9  TYPE: PRT
10 ORGANISM: Homo sapiens
11 US-09-935-868-8

```

Query Match	37.3%	Score 1275;	DB 9;	Length 592;
Post Total Similarity	47.3%	Doc No 8	48.03	

```

OY 20 PATTGNKVLGKGGTVLTLCTTASQ-KKSIOFHKNNSNQIKLGNQGSFLTKGSPSLND 78
Db 26 PNOEVARGVLTSLPGDSVLTLCPGVEPREDNAVHM-----VLKRFDA----- 66
OY 79 RADSRRLWDQGNFLLIKNLKIEDSDTYICE-----VEDQKEVOLLVFLGT 126
Db 67 -AGSHPSRWAGMGRRLLILRSVQLHDSGNVSCYRAGRAPAGTVHLLVDVPEEPQLSCFRKS 125
OY 127 ANSDTHLLQGGSLTLTLSPSSSPSVOCRSRGKNIQGGKTLISVQLELDSDSTWTCV 186
Db 126 PLSN-----VVCWGPSPSTSLTTKA-----VLLVRFKQNSPADFOEPC 165
OY 187 LQNKQVEFKIDIVVLAFOKASSIYVKKGEQVEFPFLATFVEKLNGSGEL----- 238
Db 166 QYSQESQKRSQCLAVPEGSDSFYIYSMCVASSVSGSKFTQTPQ---GGGILQDPDPANI 222
OY 239 -----W-----WQERASSSKSM-ITPDLKNKEVSVKRVTPQDPKLQMGKKLPHLHT 283
Db 223 TVTAVARNPRMISVTMODPHSMWNSFFYRLRELFELRYARASRXTFT---TMMVXDLQHHCV 278
OY 284 LPQALPQVAGSGLTLALAKTGXKHLQEOVLNVMKATQLOKNLICEVWGPTSPRLMLSLK 343
Db 279 IH-----DAMSGLRH---VQOLRA---QEEFQGGESKSEKSPANMTPW 315
OY 344 LENK----EAKVSKREKPPVVLNPEAGMWQCLLSDGVLLESNIKVLPTWSPTVPCCAP 359
Db 316 TESRSPPAENEVS--TPMQALTTKKDDNLIIFRDS-----ANATSLPVDQ-----AG 360
OY 400 EPKSCDKTKTC-----PELLGSPSYFLPPPKKDTLMTSKRPETVCVVDVSHEDPEVK 454
Db 361 EPKSCDKTKTCPCPAPELLGSPSYFLPPPKKDTLMTSKRPETVCVVDVSHEDPEVK 420
OY 455 NMVVDQVEVHNNAKTKREEQVNSTYRVVSVLTVLHQDLNKEKEKCKVSNALPAPIEKT 514
Db 421 NMVVDQVEVHNNAKTKREEQVNSTYRVVSVLTVLHQDLNKEKEKCKVSNALPAPIEKT 480
OY 515 ISKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGYPSDIAVEMESNQGPENNYKTPP 574
Db 481 ISKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGYPSDIAVEMESNQGPENNYKTPP 540
OY 575 PVLDSQGSFPLYSKLTVDKSRMQQGNVSCSMYHEALNNHTQKSLSLSPG 625
Db 541 PVLDSQGSFPLYSKLTVDKSRMQQGNVSCSMYHEALNNHTQKSLSLSPG 591

```

RESULT 30
US-10-287-035-8
; Sequence 8, Application US/10287035
; Publication No. US20030104567A1
; GENERAL INFORMATION:
; APPLICANT: Neil Stahl and George D. Yancopoulos
; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING
; FILE REFERENCE: REG 2013A
; CURRENT APPLICATION NUMBER: US/10/287,035
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: USSN 09/935,868

```

? PRIOR FILING DATE: 2001-08-23
? PRIOR APPLICATION NUMBER: USSN 09/787,835
? PRIOR FILING DATE: 2001-03-22
? PRIOR APPLICATION NUMBER: USSN 09/313,942
? PRIOR FILING DATE: 1999-05-19
? PRIOR APPLICATION NUMBER: 09/313,942
? PRIOR FILING DATE: 1999-05-19
? PRIOR APPLICATION NUMBER: 60/101,858
? PRIOR FILING DATE: 1998-09-25
? NUMBER OF SEQ ID NOS: 60
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO: 8
? LENGTH: 592
? TYPE: DRT
? ORGANISM: Homo sapiens
US-10-287-035-8

```

Query Match	37.3%;	Score 1275;	DB 14;	Length 592
-------------	--------	-------------	--------	------------

Qy	20	PAALQGNKVVLGGKKDQVVELTCTASQ-KKSLQFHKMNSNQJLILGNQSSFLTKGSSKIND	78
Db	26	PAGCVAGAVGLSLRBDGSVTLTCPCVEEBEDNATVYM-----VLARKA-----	66
Qy	79	RADSRSLWDQGNFLLIKLIKIEDSDTYICE-----VEDQKEVQVLLVFGLT	126
Db	67	-AGSHPRMAGMGRRLLRLRSYQTLDSNGNYSQCYRAGRPAQTWLLVDVPEEPQSLSCFRKS	125
Qy	127	ANSTHLLQOQSULTLTLESPPGSSPSVQCSFRKKNIOGGKTLISVQLELDQSDGTWCTCV	186
Db	126	PLSN-----VVCWGGPRSTPSLTYKA-----VLARKQNSPAEDFOEBC	165
Qy	187	LONQKVEFKIDIVLAFORASSIVYKKEGQVEFFSPFLAETVEKLTGSGEL-----	238
Db	166	QYGESEQKFCGQQLAVPBGDSSFYIVSMCVASSVSGSKFSKTQFQ---GGGILQRPANI	222
Qy	239	-----W-----WQAEPASSKSM-ITFDLKNKEVSVKRVOTDPLQNGKKLPLHIT	283
Db	223	TVTAVANPRSLVTWQDPHSMWSSFYRLRPELRYRARSKTFT---TWMTKDQHHCV	278
Qy	284	LPOALPQYAGSGNLTLEAKTGTLHQEVLVWVRATQLOKNLTCCEVWGPTSPKXLSLJK	343
Db	279	IH-----DAMSGLRH-----VVOJRA---QEEFQGEHSEMSPEAMGTPW	315
Qy	344	LENK-----EAKVSKREKRVWVLNBEAGMOCILSDSGVLLSESNIKVLPTWSTPVCAP	399
Db	316	TESRSPAEWNEVS---TPMQALTTNNKDDNILLFRDS-----ANATSLPVQD-----AG	360
Qy	400	EPKSCDKTHTC-----PELLGSEVFLPPRPKQKTLMSRTPEYTCVVVDVSHEDPEYKF	454
Db	361	EPKSCDKTHTCPPCPAPELLGGPVSFLPPRPKQKTLMSRTPEYTCVVVDVSHEDPEYKF	420
Qy	455	NMYVDGVEVNAKTKPREEQYNSTRYVVSVLTVLHODMLNKEKYCKVSNKALPAPIEKT	514
Db	421	NMYVDGVEVNAKTKPREEQYNSTRYVVSVLTVLHODMLNKEKYCKVSNKALPAPIEKT	480
Qy	515	ISKAKGQPREQVYTLTPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGQEPENNYKTPP	574
Db	481	ISKAKGQPREQVYTLTPPSRDELTKNQSLTCLVKGFYPSDIAVEMESNGQEPENNYKTPP	540
Qy	575	PVLDSDSFPLYSKLTIVDKSRWQOQGNFSCGVMEHALNNHTQKSLSLSPG	625
Db	541	PVLDSDSFPLYSKLTIVDKSRWQOQGNFSCGVMEHALNNHTQKSLSLSPG	591

RESULT 31
US-10-282-162-8
Sequence 8, Application US/10282162
Publication NO. US20030143697A1
GENERAL INFORMATION:
APPLICANT: REGENERON PHARMACEUTICALS, INC.
TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS,
AND METHODS OF MAKING

```

; TITLE OF INVENTION: AND USING
; FILE REFERENCE: REG 203-B-US
; CURRENT APPLICATION NUMBER: US/10/282,162
; CURRENT FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: 09/787,835
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 592
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-282-162-8

```

```

Query Match      37.3%; Score 1275; DB 14; Length 592;
Best Local Similarity 47.3%; Pred. No. 8,4e-82;
Matches 308; Conservative 53; Mismatches 160; Indels 130; Gaps 19;

```

```

QY 20 PAATGKNVYLGGKGDVLELTCTASQ-KKSIQFMKNSNQIKLGNQSFILTKGPKSLND 78
DB 26 PAQAVARGLVLSLPEDSVTLTCPCVGEPEEDNATVM-----VLRKPA----- 66
QY 79 RADSRRLMDGNPFLIIKNLIKIEDSDTYICE-----VEDQKEEVLIVFGLT 126
DB 67 -AGSHPRWAGMGRLLRSVQYLDHSGNYSCYRAGRAGTGHLLVDVPEEPQLSCFRKS 125
QY 127 ANSDPHLLOGSLTLTLESPPGSSPVOCSPBCKNIQOGKTLVSQLELDQSGTWCTV 186
DB 126 PLSN-----VVCWGPSTSLTTKA-----VLLVRKFQNSPAEDFQPC 165
QY 187 LONQKVEFKIDIVLAFQKASIVYKKKEGVEFSPFLAFVETKLTSSGL----- 238
DB 166 QYSGESQKFSQCLAVPEGDSSFYIVSMCVASVSGSKSTQTFQ---GCGILQDPDPANI 222
QY 239 -----WQAEKASSSKSM-ITFDLKNKEVSVRKVTQDPKLOMGKTLPLTLT 283
DB 223 TVTAVARNPRLSVTWODPHSNSSFYRLRFELRYABRSKTF-----TMMVKDLQHNVCV 278
QY 284 LPQALPOYAGSGNLTLEAKTGKLDQEVNLVWMRATOLQKLTCEWGPRTSPKLMISLK 343
DB 279 IH-----DAWSGLRH-----VVOJRA---QEFQGGSEWSEWSPRAMGTPW 315
QY 344 LENK-----EAKVSKREKPVWNLNPEAGMOCILSDSQVLESNIKVLPTMTSTVPCAP 399
DB 316 TESRRPRAENEVS---TPWQALTTNKDDNLTFRDS-----ANATSLPVQD-----AG 360
QY 400 EPKSCDKTHTC-----PELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEYKF 454
DB 361 EPKSCDKTHTCPCAPPELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEYKF 420
QY 455 NMVYDGVENNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYKCKVSKALPAPIEKT 514
DB 421 NMVYDGVENNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYKCKVSKALPAPIEKT 480
QY 515 ISKAKGOREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEMBSNQPENNYKTPR 574
DB 481 ISKAKGOREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEMBSNQPENNYKTPR 540
QY 575 PVLSDSGSFLYSKLTVDKSRMOQGNVFSQVMHEALAHNYTOKSLSLSPG 625
DB 541 PVLSDSGSFLYSKLTVDKSRMOQGNVFSQVMHEALAHNYTOKSLSLSPG 591

```

```

RESULT 32
US-09-815-108-22
; Sequence 22, Application US/09815108
; Patent No. US20020009776A1
; GENERAL INFORMATION:
; APPLICANT: Sarris, Christiaan M.
; APPLICANT: Sharon, Wu X.
; APPLICANT: Xia, Min

```

```

; APPLICANT: Boone, Thomas Charles
; APPLICANT: Covey, Todd
; TITLE OF INVENTION: Fibroblast Growth Factor Receptor-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 99-513-A
; CURRENT APPLICATION NUMBER: US/09/815,108
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/191,379
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: murine FGFR-L
; OTHER INFORMATION: extracellular domain-Fc fusion polypeptide
US-09-815-108-22

```

```

Query Match      37.1%; Score 1265; DB 9; Length 594;
Best Local Similarity 44.2%; Pred. No. 4.3e-81;
Matches 311; Conservative 55; Mismatches 132; Indels 206; Gaps 22;

```

```

QY 11 LVLVQLALPLPA-----TQGNKV-----LQK-----KGDVLELTCTASQKSI 49
DB 7 LLLILLALSLAEARPPRMADKVPYVARLGRVRLQCPYEGDPPPLTMMTKGRTI 66
QY 50 QFMKNSNQIKLGNQSFILTKGPKSLNDRADSRRLMDGNPFLIIKNLIKIEDSDTYIC 109
DB 67 HSGW-----SRFVLPQG-----LKVEAEADAGVVC 95
QY 110 EVDQKEEVLIVFGLTANSDTHLQGSLLTLTLESP-PESS----- 150
DB 96 KATNG-----FGSLSVYTLTIM-DDISPKESPGSGSSGGQEDPASQGWARPFT 146
QY 151 -PSVQCR-----SPRKNIOGSK-----TLSVSQ 173
DB 147 QPSGMRRLVRLAPRGSSVRLKCVASGHPRPDIIMMKDDQTLTHLSEHKKKKTLSLN 206
QY 174 LEHDQSGTWCTVYLQNKQVE--FKIDIVLAFQKASIVYKKKEGVEFSPFLAFVETK 231
DB 207 LKPEDSGKYTCRVSNKAGAINATYKVDVIORTSKPVLGTGHPVNTTVDGCTTSFOCK- 265
QY 232 LTGSGELMWAERASSSKSIITFDLKNKEVSVRKVTQDPKLOMGKTLPLTLPOALPOY 291
DB 266 -----VRSDVK-----PVIQWLKRV-----EY 282
QY 292 AGSGNLTLEAKTGKLDQEVNLVWMRATOLQKLTCEWV-----GPTSPKLMISLKLENK 347
DB 283 GSEGRHNSTIDVGQK-----FVULP-----TGVMSRPDGSYLNKLLISRAQD- 327
QY 348 EAKVSKREKPVWNLNPEAGMOCILSDS--GQVLESNIKVLPTMTSTVPCAPPEPKSCDK 406
DB 328 -----DAGNYICIGANTMGYSFRSAFLTVLPDPPKPGPMA--SSSVDK 369
QY 407 THTC-----PELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEYKVMYVDGV 461
DB 370 THTCPCPAPPELLGSPSVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEYKVMYVDGV 429
QY 462 EVNNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYKCKVSKALPAPIEKTISAKQ 521
DB 430 EVNNAKTKPREEOYNSTYRVSVLTVLHODMLNGKEYKCKVSKALPAPIEKTISAKQ 489
QY 522 PREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEMBSNQPENNYKTPRVLSDSG 581
DB 490 PREPOVYTLPPSRDELTKNOVSLTCLVKGFPDIAVEMBSNQPENNYKTPRVLSDSG 549
QY 582 SFFLYSKLTVDKSRMOQGNVFSQVMHEALAHNYTOKSLSLSPG 625
DB 550 SFFLYSKLTVDKSRMOQGNVFSQVMHEALAHNYTOKSLSLSPG 593

```

RESULT 33
US-10-229-584-22
; Sequence 22, Application US/10229584
; Publication No. US2003087384A1
; GENERAL INFORMATION:
; APPLICANT: Saries, Christiaan M.
; APPLICANT: Sharon, Mu X.
; APPLICANT: Xia, Min
; TITLE OF INVENTION: Fibroblast Growth Factor Receptor-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 99-513-F
; CURRENT APPLICATION NUMBER: US/10/229,584
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/815,108
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/191,379
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: murine FGFR-L
; OTHER INFORMATION: extracellular domain-Fc fusion polypeptide
US-10-229-584-22

Query Match 37.1%; Score 1265; DB 14; Length 594;
Best Local Similarity 44.2%; Pred. No. 4.3e-81;
Matches 311; Conservative 55; Mismatches 132; Indels 206; Gaps 22;

QY 11 LVLQLALPLA-----TQGNKV-----LQK-----KQDVELTCAQSKSI 49
DB 7 LLLLLALPSAEARPPPMADKVPQYARLGRVRLQCPVGEPPPLTWTAKGRTI 66
QY 50 QPFWKNSNQIKILGNQGSFLTKGPSKLNDRASRSRLMDQGNFPLIKIKIIDSPTIYC 109
DB 67 HSGH-----SRRFLPQG-----LKKVEADAGAYVC 95
QY 110 EVEDQKEEVLVFGLTANSDFLLQGSITLTLESF-PGSS-----TLVSQ 150
DB 96 KATNG-----FGSLVNYVTLIM-DisPKESPQSGSGQGPASQOMABRFT 146
QY 151 -PSVQCR-----SPRKNIOGK-----TLVSQ 173
DB 147 QPSPMRRTVIAIPVSSVRLKCVASGHPRPDIIMMKDQTLTLLEASERKKKTUSLKN 206
QY 174 LEIADSGTWTCTVLQONOKVE--FKIDIVLAFQKASSIYVKKEGEVEFSFPLAFTVEK 231
DB 207 LKPEDSGKYTCRYSNKGALNATYKVDVIGRTSKPVLGTHTPVNTTVDGCTTSFOCK- 265
QY 232 LTGSGELMWAERASSKSWITPDLKNEKVSVKRVTDPRKQNGKKPLHLTLPLQALPOY 291
DB 266 -----VRSBVK-----PVIQMLKRV-----EX 282
QY 292 AGSGLNLLEAKTKGLHGVNLVYMAATQLOKNTLGEVW-----GPTSPKMLSLKLENK 347
DB 283 GSEGRHNSITDVGQK-----FVILP-----TDGWSRPDGSYLNKLLSRAQD- 327
QY 348 EAKVSKREKPVWVLPNPEAGMOCCLSDS-GVLLLESNIKYLPTWSTVPAPAPESKCDK 406
DB 328 -----DAGMYICIGANTMGYSRPSALVTLVLPDKPGPMA-----SSVDK 369
QY 407 THTC-----PELLGSPVFLPFPKPKDTLMI SRTPEVTCVAVDVSHDEPEVKFNWYDGV 461
DB 370 THTCPAPAPELLGSPVFLPFPKPKDTLMI SRTPEVTCVAVDVSHDEPEVKFNWYDGV 429
QY 462 EVNNAKTRPREQVNSTYRVVSVLTVLHOMLNKKEKCKVSNKALPAPLEKTSKAKGQ 521
DB 430 EVNNAKTRPREQVNSTYRVVSVLTVLHOMLNKKEKCKVSNKALPAPLEKTSKAKGQ 489
QY 522 PREQVNTLPSRDELTKNQVSLTCLVKGYFSPDIAVEMESNGQPENNYKTTTPVLSDG 581

DB 490 PREQVNTLPSRDELTKNQVSLTCLVKGYFSPDIAVEMESNGQPENNYKTTTPVLSDG 549
QY 582 SFPLYSKLTVDKSRMQQGNVFCSSVMHEALNHHYTKQSLISLSPG 625
DB 550 SFPLYSKLTVDKSRMQQGNVFCSSVMHEALNHHYTKQSLISLSPG 593

RESULT 34
US-09-910-600-16
; Sequence 16, Application US/09910600
; Publication No. US20030036631A1
; GENERAL INFORMATION:
; APPLICANT: Longphre, Malinda
; APPLICANT: Chang, Han
; APPLICANT: Whitley, Gena
; TITLE OF INVENTION: NOVEL SIGLECS AND USES THEREOF
; FILE REFERENCE: D0003NP
; CURRENT APPLICATION NUMBER: US/09/910,600
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/220,139
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: L3-hlg
US-09-910-600-16

Query Match 37.1%; Score 1265; DB 10; Length 779;
Best Local Similarity 45.0%; Pred. No. 6.2e-81;
Matches 304; Conservative 43; Mismatches 136; Indels 192; Gaps 19;
QY 8 RHLLVLQALPLA--ATQGNKVVL-GKKQDVELTCAQSKSIQPFMKNSNQIKILG 63
DB 239 RDLVISISNDTPALPEPQQGNVPLYLQKQGFLLCLADQGPRTLSW----- 288
QY 64 NQGSFLTKGPSKLNDRASRSRLMDQGNFPLIKIKIIDSPTIYCEVEDQ----- 114
DB 289 -----VLQNRVLSSSHMPQRPPLGLPLPGYKAGDSGRYTCRAENRLGSGORALD 337
QY 115 -----KEEVOLVLF-----GLTANSDFLLQGSITLTLESFPGSSPSVQCS 157
DB 338 LSVQYPPENLVRVWVQANFTVLNGLNGTSLPVLGQSCLVGVTHSSPPA----- 388
QY 158 PRGKNIOGKTSVSO-----LEIADSGTWTCTVLQONOKVEFKIDIVLAFQK 206
DB 389 -RLSWQKQVLSPEQSPDPGVLELPRVQVHEGEFTCHAR----- 428
QY 207 ASSIYVKKEGEVEFSFPLAFTVEKLTGSGELW-----WQNERASSKSWITPDL 256
DB 429 -----HPLGSGVHVSLSLVHYS-PKLLGPSCSWEAEGILHCSSQSPAPSLRWML- 478
QY 257 KNEKVSVKRVTDPRKQNGKKPLHLTLPLQALPOYAGS-GNITLLEAKTKGLHGVNLV 315
DB 429 -----HPLGSGVHVSLSLVHYS-PKLLGPSCSWEAEGILHCSSQSPAPSLRWML- 478
QY 479 -GEELLEGNSQDSF-----EVTSSAGPWANSSLSLH-----GGLSGL- 517
DB 316 VNRATQLOKNTLCEVWGPTSPKMLSLKLENKAKVSKREKPVWVLPNPEAGMOCCLSDS 375
QY 518 -----RLRCEANNVGAQSGSLQJLPDKG-----LISD- 546
DB 376 GOVLLLESNIKYLPTWSTVPAPAPESKCDKTHTC-----PELLGSPVFLPFPKPKDTL 430
QY 547 -----PEPKSCDKTHTCPPCAPAPPEGAPSVFLPFPKPKDTL 583
QY 431 MTSRTEVTCVAVDVSHDEPEVKFNWYDGVVNAKTRPREQVNSTYRVVSVLTVLHQ 490
DB 584 MTSRTEVTCVAVDVSHDEPEVKFNWYDGVVNAKTRPREQVNSTYRVVSVLTVLHQ 643
QY 491 DMLNGREYKCKVSNKALPAPLEKTSKAKGQPREQVNTLPSRDELTKNQVSLTCLVKG 550

Db 644 DWLNGKEYKCKVSKNALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKG 703
Qy 551 FYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSCVWHEA 610
Db 704 FYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSCVWHEA 763
Qy 611 LHHNYTKSLSLSPG 625
Db 764 LHHNYTKSLSLSPG 778

RESULT 35

US-09-910-600-30
; Sequence 30, Application US/09910600
; Publication No. US20030036631A1
; GENERAL INFORMATION:
; APPLICANT: Longphre, Malinda
; APPLICANT: Chang, Han
; APPLICANT: Miltney, Gena
; TITLE OF INVENTION: NOVEL SICLECS AND USES THEREOF
; FILE REFERENCE: D0003ND
; CURRENT APPLICATION NUMBER: US/09/910,600
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/220,139
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: L3-hig
US-09-910-600-30

Query Match 37.1%; Score 1265; DB 10; Length 779;
Best Local Similarity 45.0%; Pred. No. 6.2e-81;
Matches 304; Conservative 43; Mismatches 136; Indels 192; Gaps 19;
Qy 8 RHLLVLQALLPA--ATQGNKVVLL-GKKGDVLELTCTASOKKSIOFHWKNSNOIKILG 63
Db 239 RDLVISISRDNTPALPEPQGNVPLBAQKQFLRLCAADSOPTATLSW----- 288
Qy 64 NQGSFLTGSPSKLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQ----- 114
Db 289 -----VLONRRVLSSSHPMGPRPLGELPEGVKAGDSGRYTCRAENLIGSQORALD 337
Qy 115 -----KEVOLVLF-----GLTANSDTHLLOGOSLTL--TLESPPGSSPSVQCORS 157
Db 338 LSVQYPPENLKVMSQANRYLENLNGTSLPVEGOSLCLVCTHTSSPPA----- 388
Qy 158 PRGNKIQGKTLVSQO-----LELDOSGTWCTVLONOKKVEFKIDIVLVAFOK 206
Db 389 -RLSWGQGVLSPSQPSDPGVLELPRVQVEHBEFTCHAR----- 428
Qy 207 ASSIYVKKGEQVESPLAFTVEKLTGSGELM-----MOAERASSKSWITFDL 256
Db 429 -----HPLGSHVSLSLSVHYS-PKLLGPSCSWEABGHLCCSSQASAPSLRWML---- 478
Qy 257 KKEKESVVRVTPQDPKLQWKKLPLHLTLPOALPOYAGS-GNLTALAEKTKLQOEVLV 315
Db 479 -GEBLEBNSQDSF-----EVTPESSAGMANSSLSLH---GGLSSGL-- 517
Qy 316 VWRATOLQKNTCEVWGPTSPKMLSLKLENKEAKVSKREKRVVWVLANPEAGMOCLLSDS 375
Db 518 -----RLRCEAMNVHGAQSGSIQLPDKG-----LISD- 546
Qy 376 GOVLLSNIKVLPTWSTVPCCAPAPKSCDKTHTC-----PELLGSPVFLFPKPKDTL 430
Db 547 -----PEPKSCDKTHTCPCCAPAPFPKAPSVFLFPKPKDTL 583
Qy 431 MISRTEVTCVVVDVSHEDPEVKFMWYVDGVEVNAKTKPREEQNGNSTRVAVSLTVLHQ 490

Db 584 MISRTEVTCVVVDVSHEDPEVKFMWYVDGVEVNAKTKPREEQNGNSTRVAVSLTVLHQ 643
Qy 491 DWLNGKEYKCKVSKNALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKG 550
Db 644 DWLNGKEYKCKVSKNALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKG 703
Qy 551 FYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSCVWHEA 610
Db 704 FYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSCVWHEA 763
Qy 611 LHHNYTKSLSLSPG 625
Db 764 LHHNYTKSLSLSPG 778

RESULT 36

US-10-207-655-348
; Sequence 348, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069,401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 348
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-348

Query Match 37.0%; Score 1262; DB 14; Length 504;
Best Local Similarity 46.7%; Pred. No. 5.7e-81;
Matches 302; Conservative 42; Mismatches 120; Indels 182; Gaps 20;
Qy 1 MNRGVPRRHLLVQLALPAAQGNKVVLLGKKGDVLELTCTASOKKSIOFHW--KNSN 57
Db 19 MSRGVDIQ-----MTQTSSLASIGDRRTICRASQDIRNLYNWYQKPDG 65
Qy 58 QIKILNQGSFPLTKGPKLNDRADSRSLMDQGNFPLIKLKIEDSDTYICEVEDQKE 116
Db 66 TVKXL-----IYYT---SRLHSGVPSRFGSGSGTDYSLTIANLQPEDIAITYFCQ----- 112
Qy 117 EVQLVLRGTLANSDTHLLOGOSLTLTLESPPGSSPSVQCORSPPGKNIQGGKT---LSVSQ 173
Db 113 -----QNTLPLWTF---GGGTVKLTVKELQGGGGGGGGGGSIDE 151
Qy 174 LELDOSGTWCTVLONOKKVEFKIDIVLVAFOKASSIVYKKEGEQVESPLAFTVEKLT 233
Db 152 VOLOQSGP-----ELV-----KQASMSCAASG---YSF-TGIYV--- 183
Qy 234 GSGELMWQAEARASSKSWITFDLKNKEVSVKRVTPQDKLOMGKKLPLHLTLPOALPOYAG 293
Db 184 -----WLKQSHGKNLEWIGLINPYKGLT---TYNQKFK----- 213
Qy 294 SGNLTLLAEKTKLQOEVLVWVRATOLQKNTLCEVWGPTSPKMLSLKLENKEAKVSK 353
Db 214 -GRATLTVDKSSSTAYVE-----LISLTSSEDSAVYYCA 245
Qy 354 REK-----PYVVLNPEAGMOCLLSDSGVLLSNIKVLPTWSTVPCCAPAPKSC 404
Db 246 RSGYIGSDMYFDVW---GAGTTVYSSIQ-----EKSK 277
Qy 405 DKHTTC-----PELLGSPVFLFPKPKDTLMSRTEVTCVVVDVSHEDPEVKFMWYVD 459
Db 278 DKHTTCPPCAPAPPELLGSPVFLFPKPKDTLMSRTEVTCVVVDVSHEDPEVKFMWYVD 337

```

QY 460 GVEHNAKTPREBOYNSTYRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAK 519
DB 338 GVEHNAKTPREBOYNSTYRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAK 397
QY 520 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG 579
DB 398 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG 457
QY 580 DGSFFLYSKLTVDKSRWQGQNVFSCVMHEALHNHYTQKSLSLSPG 625
DB 458 DGSFFLYSKLTVDKSRWQGQNVFSCVMHEALHNHYTQKSLSLSPG 503

RESULT 37
US-10-207-655-345
; Sequence 345, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207.655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 345
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-345

Query Match 36.9%; Score 1260; DB 14; Length 543;
Best Local Similarity 46.4%; Pred. No. 8.7e-81;
Matches 299; Conservative 43; Mismatches 112; Indels 190; Gaps 19;

QY 1 MNRGVPFRHLILVLOLALPAATQGNKVVLGKGDVTELTCTASOKKSIOFHMKNSTQIK 60
DB 19 MSRGVD-----IVL-----TQSPPTTAAISPGKEVTITCRASSSVSYVMWYQOKS--- 62
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQG-NPFLIIKLNKIEDSDTYICEVEDQKEBQV 119
DB 63 --GASPKLMIYDTSKLAGVPNRFSGSGSTYSIAINMETEDATATYCO----- 111
QY 120 LLVFGLTANSDDLHLOGQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLTSLVSOLELDS 179
DB 112 -----QWSTPLTF-----GSQTKLEIKRGGSGGSGGSGGQVQLKKA 152
QY 180 G-----TWCTVQLQNKQKVEFKIDIVLAFQKASSIYVKKEGEQVESEFPLAFTV 229
DB 153 GPGIVQPTQTLSLTCTV-----SGFS----- 173
QY 230 EKLTSGSELWMQAEBAASSKSWITFDLKNKEVSVKRVTDODPKLQNGKKLPMLHLTPQALP 289
DB 174 --LTSQGVH-----IRQPP-----GKGLEW-----MGII 196
QY 290 QYAGSGNLTLAEAKTKGLHGEVNLVVMRATQLOKNLTCEWGPSTPKMLSLKENKEA 349
DB 197 YIDGGTDYNSAIKSR-----LSTSRDTS-----KQGVFLKINSLO- 231
QY 350 KVSKEKEPWWLNPQAGMOCQ---LSDSGVLLSNIKVLPTWSTVPVCPAPPEPSCDK 406
DB 232 -----TDDTMYYCARIHFDWQO-----GVMVTVSSDL-----EPKSCDK 267
QY 407 THTC-----PELLGGSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 461
DB 268 THTCPCPAPBELLGGSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 327
QY 462 EVHNAKTPREBOYNSTYRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQ 521
DB 328 EVHNAKTPREBOYNSTYRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQ 387

```

```

QY 522 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG 581
DB 388 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG 447
QY 582 SFFLYSKLTVDKSRWQGQNVFSCVMHEALHNHYTQKSLSLSPG 625
DB 448 SFFLYSKLTVDKSRWQGQNVFSCVMHEALHNHYTQKSLSLSPG 491

RESULT 38
US-10-207-655-344
; Sequence 344, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207.655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 344
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-344

Query Match 36.8%; Score 1258; DB 14; Length 492;
Best Local Similarity 46.4%; Pred. No. 1.1e-80;
Matches 299; Conservative 42; Mismatches 113; Indels 190; Gaps 19;

```

```

QY 1 MNRGVPFRHLILVLOLALPAATQGNKVVLGKGDVTELTCTASOKKSIOFHMKNSTQIK 60
DB 19 MSRGVD-----IVL-----TQSPPTTAAISPGKEVTITCRASSSVSYVMWYQOKS--- 62
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQG-NPFLIIKLNKIEDSDTYICEVEDQKEBQV 119
DB 63 --GASPKLMIYDTSKLAGVPNRFSGSGSTYSIAINMETEDATATYCO----- 111
QY 120 LLVFGLTANSDDLHLOGQSLTLTLESPGSSPVQCRSPRGKNIQGGKTLTSLVSOLELDS 179
DB 112 -----QWSTPLTF-----GSQTKLEIKRGGSGGSGGSGGQVQLKKA 152
QY 180 G-----TWCTVQLQNKQKVEFKIDIVLAFQKASSIYVKKEGEQVESEFPLAFTV 229
DB 153 GPGIVQPTQTLSLTCTV-----SGFS----- 173
QY 230 EKLTSGSELWMQAEBAASSKSWITFDLKNKEVSVKRVTDODPKLQNGKKLPMLHLTPQALP 289
DB 174 --LTSQGVH-----IRQPP-----GKGLEW-----MGII 196
QY 290 QYAGSGNLTLAEAKTKGLHGEVNLVVMRATQLOKNLTCEWGPSTPKMLSLKENKEA 349
DB 197 YIDGGTDYNSAIKSR-----LSTSRDTS-----KQGVFLKINSLO- 231
QY 350 KVSKEKEPWWLNPQAGMOCQ---LSDSGVLLSNIKVLPTWSTVPVCPAPPEPSCDK 406
DB 232 -----TDDTMYYCARIHFDWQO-----GVMVTVSSDL-----EPKSCDK 267
QY 407 THTC-----PELLGGSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 461
DB 268 THTCPCPAPBELLGGSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 327
QY 462 EVHNAKTPREBOYNSTYRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQ 521
DB 328 EVHNAKTPREBOYNSTYRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQ 387
QY 522 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDG 581
DB 328 EVHNAKTPREBOYNSTYRVVSVLTVLHODMLNGEKYCKVSNKALPAPIEKTISKAKGQ 387

```

Db 388 PREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDG 447
Qy 582 SFFLYSKLTVDKSRMOQGNVFCSCVMEALHNHYTQKSLSLSPG 625
Db 448 SFFLYSKLTVDKSRMOQGNVFCSCVMEALHNHYTQKSLSLSPG 491

RESULT 39
US-09-773-877A-12
; Sequence 12, Application US/09773877A
; Publication No. US20030017977A1
; GENERAL INFORMATION:
; APPLICANT: Xia, Yu-Ping et al.
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY SKIN DISEASES
; FILE REFERENCE: REG 710b
; CURRENT APPLICATION NUMBER: US/09/773,877A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 567
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flt(1-3)-Fc
US-09-773-877A-12

Query Match 36.7%; Score 1254.5; DB 12; Length 567;
Best Local Similarity 47.2%; Pred. No. 2.3e-80;
Matches 298; Conservative 55; Mismatches 120; Indels 159; Gaps 24;

Qy 85 SLMDQGNFPLIIKLIKIEDSDTYICEVDOKEEVQLVFGLTANSDFHLQ-GGSLTTLT 143
Db 3 SYMDTVGLCALISCLILTGSSSGSKLDPE-----LSLKGTHIMQAGOTLHL-- 51
Qy 144 ESPPGSSPSVOCR-----SPRGKNIQGRK---TLVSQLE 175
Db 52 -----QCRGEAAHKMSLPEWVSKESERLSITSACGRN---GKQFCSTLTLNTAQ 98
Qy 176 LODSGTWTCTVL---ONOKKVERKIDI-----VTLAFQKASSIVYKKEGEQVERFSF 223
Db 99 ANHGFYSCKYLAIVPTSKKETESAIYFIISDTGRPFVEMWSEIPEIIMHTEGR--ELVI 156
Qy 224 PLATVEKLT-----GSGELMWQABRASSKSWITFDLKNKEVSVKRYTOD 269
Db 157 PCRTVSPNITVTLKKPFLDTLIPDGKRIW-----DSRKGFIISNATYKEIGL----- 204
Qy 270 PKLQMGKKLPLHLTLPLQALPOYAGSGNLTALBAKTGKL--HQEVNLVY-----MRATOL 322
Db 205 -----LTCBATV-----NGHL-----YKTNVLTNRQNTIITDVOISTPRPVKL 242
Qy 323 QKN---LTCBVWGPTSPKMLSLKL--ENKEAKVSKR-----EKPV 358
Db 243 LRGHTLVNCTATTPLNTRVQMTWSYPDEKKNKASVRRRIDQSNSHANIFYSVLTIDK-- 300
Qy 359 WVLNPEAGMOCCLSDSGQVLLSESNIKVLPWTSTPVPAPAPKSCDKTHTC-----PEL 413
Db 301 -MOKKDKGLTYCRRV--SGPSFKSVNTSV--HIYDAGP---GEPKSCDKTHTCPCPAPBL 354
Qy 414 LGGSPVLFPPKPKDPTLMISRTPEVTCVVDVSHEDDEVKFNMYVVDGEVHNATKTRRE 473
Db 355 LGGSPVLFPPKPKDPTLMISRTPEVTCVVDVSHEDDEVKFNMYVVDGEVHNATKTRRE 414
Qy 474 QYNSTYRVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISAKAQPREPQVYTLPPS 533
Db 415 QYNSTYRVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISAKAQPREPQVYTLPPS 474
Qy 534 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK 593
Db 475 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK 534
Qy 594 SRMOQGNVFCSCVMEALHNHYTQKSLSLSPG 625
Db 594 SRMOQGNVFCSCVMEALHNHYTQKSLSLSPG 625

Db 535 SRMOQGNVFCSCVMEALHNHYTQKSLSLSPG 566

RESULT 40
US-09-773-877A-20
; Sequence 20, Application US/09773877A
; Publication No. US20030017977A1
; GENERAL INFORMATION:
; APPLICANT: Xia, Yu-Ping et al.
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY SKIN DISEASES
; FILE REFERENCE: REG 710b
; CURRENT APPLICATION NUMBER: US/09/773,877A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 567
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flt(1-3 R->N)-Fc (Mult4)
US-09-773-877A-20

Query Match 36.7%; Score 1254.5; DB 12; Length 567;
Best Local Similarity 47.2%; Pred. No. 2.3e-80;
Matches 298; Conservative 55; Mismatches 120; Indels 159; Gaps 24;

Qy 85 SLMDQGNFPLIIKLIKIEDSDTYICEVDOKEEVQLVFGLTANSDFHLQ-GGSLTTLT 143
Db 3 SYMDTVGLCALISCLILTGSSSGSKLDPE-----LSLKGTHIMQAGOTLHL-- 51
Qy 144 ESPPGSSPSVOCR-----SPRGKNIQGRK---TLVSQLE 175
Db 52 -----QCRGEAAHKMSLPEWVSKESERLSITSACGRN---GKQFCSTLTLNTAQ 98
Qy 176 LODSGTWTCTVL---ONOKKVERKIDI-----VTLAFQKASSIVYKKEGEQVERFSF 223
Db 99 ANHGFYSCKYLAIVPTSKKETESAIYFIISDTGRPFVEMWSEIPEIIMHTEGR--ELVI 156
Qy 224 PLATVEKLT-----GSGELMWQABRASSKSWITFDLKNKEVSVKRYTOD 269
Db 157 PCRTVSPNITVTLKKPFLDTLIPDGKRIW-----DSRKGFIISNATYKEIGL----- 204
Qy 270 PKLQMGKKLPLHLTLPLQALPOYAGSGNLTALBAKTGKL--HQEVNLVY-----MRATOL 322
Db 205 -----LTCBATV-----NGHL-----YKTNVLTNRQNTIITDVOISTPRPVKL 242
Qy 323 QKN---LTCBVWGPTSPKMLSLKL--ENKEAKVSKR-----EKPV 358
Db 243 LRGHTLVNCTATTPLNTRVQMTWSYPDEKKNKASVRRRIDQSNSHANIFYSVLTIDK-- 300
Qy 359 WVLNPEAGMOCCLSDSGQVLLSESNIKVLPWTSTPVPAPAPKSCDKTHTC-----PEL 413
Db 301 -MOKKDKGLTYCRRV--SGPSFKSVNTSV--HIYDAGP---GEPKSCDKTHTCPCPAPBL 354
Qy 414 LGGSPVLFPPKPKDPTLMISRTPEVTCVVDVSHEDDEVKFNMYVVDGEVHNATKTRRE 473
Db 355 LGGSPVLFPPKPKDPTLMISRTPEVTCVVDVSHEDDEVKFNMYVVDGEVHNATKTRRE 414
Qy 474 QYNSTYRVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISAKAQPREPQVYTLPPS 533
Db 415 QYNSTYRVSVLTVLHODMLNGKEYKCKVSNKALPAPIEKTISAKAQPREPQVYTLPPS 474
Qy 534 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK 593
Db 475 RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK 534
Qy 594 SRMOQGNVFCSCVMEALHNHYTQKSLSLSPG 625
Db 535 SRMOQGNVFCSCVMEALHNHYTQKSLSLSPG 566

RESULT 41

```

US-10-412-406-33
; Sequence 33, Application US/10412406
; Publication No. US20040058394A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: GARBER, Ellen
; APPLICANT: LYNE, Paul
; APPLICANT: SALDHANA, Jose W.
; TITLE OF INVENTION: HUMANIZED ANTI-LT-BETA-R ANTIBODIES
; FILE REFERENCE: BINA100CN
; CURRENT APPLICATION NUMBER: US/10/412,406
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 60/240,285
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/275,289
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/299,987
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: PCT/US01/32140
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 4852
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-412-406-33

Query Match          36.7%; Score 1253.5; DB 12; Length 4852;
Best Local Similarity 45.5%; Pred. No. 4.8e-79;
Matches 322; Conservative 53; Mismatches 180; Indels 153; Gaps 26;

QY      23  TQGNKVVLGKGGTVELTCTASQKSIQFWKNSNQIKLIGNSGFLTKPSKLNDRADS 82
      4193  TQSPSSISASVGRVITTCAGQDIXSYLSWYQOKP-----GKAPKLIYATRLAGVPS 4248
      83  RRLMDQG-NFPLIKNLIKEDSDTYC-----EVEDQKEVQLVFLGTA 127
      4249  RFGSGSGSDYTLTISLSQPEDPATIYCLHGSPMTFGGTYLEIKRYLAASVF-IPF 4307
      128  NSDTHLLOGSLTLTLSPGSSPSVQCR---SPRGKNIQ-----GKTLVSQI 174
      4308  PSEBQKSG-----TASVVCILNFPYPRAKQWKVDNALGSGNSQESVTEQ 4354
      175  ELDDSGTWTCTVQNQKVEF---KIDIVLAFQKASSIYK--KSGEYVESFPPLAFY 229
      4355  DSKDS--TYSLSSTLTLKADYEKHKYVACEVTHQGLSSPYTKSFGNCECEVOL-----V 4408
      230  EKLTG---SGELMWQERASSSSKSWTFPDLKQKEVSVKRVTDPKLQMGKLP.LHLTL 285
      4409  ESGGGLVKKPGSL-----RLSCAASGFTF---SDIYMYFROPAP---GKGLEWATIS 4455
      286  Q--ALPOYAGS--GNLTLLAEATGK.LHQEVNLVVMKATOL-----QKN--LTCEVWG 333
      4456  DGGSYTYPPSVGRFTISRDNKNSLYLQMSLSRADTAVVYCAARENGNFYFDYWGQ 4515
      334  TSPKMLSLKLENKAEKVKREKPVVYLNPEA-----GMQCLISD-----S 375
      4516  GTTVTVSS-----ASTKGPSVFPFLAPSSKSTSGGTALGLVADYEPPEVTVSWS 4566
      376  GOVLLSENIKVL-----TWSTVPVCPAP-----EPK 402
      4567  G--ALTSGVHTFPAVLOSGLYSLSSVTVYVSSSLGTQTYICNVNHNKPSNTKVDKVEPK 4624
      403  SCDKTHTC-----PELLGSPSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKRWY 457
      4625  SCCKTHCPCPAPPELLGSPSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKRWY 4684
      458  VDGVEVNAKTKPREEOYNSTRYVSVLTVLHODMLNGEKYCKVSNKALPAPIETKISK 517
      4685  VDGVEVNAKTKPREEOYNSTRYVSVLTVLHODMLNGEKYCKVSNKALPAPIETKISK 4744
      518  AKQGRPREQVYTLPPSRDELTKQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPV 577

```

```

DB      4745  AKQGRPREQVYTLPPSRDELTKQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPV 4804
QY      578  DSDGSEFFLYSKLTVDKSRMVGQNVFSCVMEALHNHYTKSISLSPG 625
      4805  DSDGSEFFLYSKLTVDKSRMVGQNVFSCVMEALHNHYTKSISLSPG 4852

RESULT 42
US-10-207-655-266
; Sequence 266, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 266
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-266

Query Match          36.7%; Score 1252.5; DB 14; Length 552;
Best Local Similarity 46.0%; Pred. No. 3e-80;
Matches 297; Conservative 39; Mismatches 126; Indels 183; Gaps 18;

QY      1  MNRGVPFRLHLLVLTALIPATQGNKVKLGKGGTVELTCTASQKSIQFWK---KNSN 57
      19  MNRGVD-----IVL-----TQSPATLSVTPGDRVSLSCASQISIDYLMHYQOKSH 65
      58  QIKILGNOSFLTKG--PSKLNDRADSRRLMDQGNFPLIKNLIKEDSDTYICEVDDQKE 116
      66  SPRLIKTVASHISISGTPSRFSGSGC-----SDTLTINSVPEPDVGITYCQ----- 112
      117  EVQLVFGLTANSDTHLLOGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLEL 176
      113  -----HGHSPWTF---GGGTLEIKRGGSGSGSGSGSGQIQI 150
      177  QDSGTWTCTVQNQKVEFKIDI VLAFOKASSIYKKEGEYVESFP---LAFT----- 228
      151  VQSGP-----ELKKPGETVRLISCKASGYAFITTMQ 181
      229  -VEKLTSGELMWQERASSSSKSWTFPDLKQKEVSVKRVTDPKLQMGKLP.LHLTL 287
      182  WQEMPGKLMK-----IGWINTPLMSAKIC-----RRLQ----- 211
      288  LPOYAGSGLTLAEATGK.LHQEVNLVVMKATOLQKNLTCEVWGFTSPKMLSLKLENK 347
      212  -----GRFASLEISANTAVLIQIS-----NLNDE-- 235
      348  EAKVSRREKPVVYLNPEAGMQCLISDSGOVLLSENIKVLPTW--STVPVCPAPPEKSCD 405
      236  -----DTATYFCVRSNGNRY---DLAFAYWGGQGLTVVSDLEPKSSD 275
      406  KTHT-----CPELLGSPSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKRWY 460
      276  KTHTSPSPAPPELLGSSVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKRWY 335
      461  VEVHNAKTKPREEOYNSTRYVSVLTVLHODMLNGEKYCKVSNKALPAPIETKISK 520
      336  VEVHNAKTKPREEOYNSTRYVSVLTVLHODMLNGEKYCKVSNKALPAPIETKISK 395
      521  QPREQVYTLPPSRDELTKQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPV 580
      396  QPREQVYTLPPSRDELTKQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPV 455

```


Db 625 SVLTVLHODWLNKCKVKNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQV 684
QY 543 SLTCLVKGFPYSDIAEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNV 602
Db 685 SLTCLVKGFPYSDIAEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNV 744
QY 603 SCSVMHEALHNHYTOKSLSPG 625
Db 745 SCSVMHEALHNHYTOKSLSPG 767

RESULT 45
US-09-875-338-9
Sequence 9, Application US/09875338
Patent No. US20020093024A1
GENERAL INFORMATION:
APPLICANT: MIKESSELL, GLEN E.
APPLICANT: CHANG, HAN
APPLICANT: FINGER, JOSHUA N.
APPLICANT: YANG, GUCHEN
APPLICANT: LU, PIN
APPLICANT: ZHOU, XIA-DI
APPLICANT: PEACH, ROBERT
TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
FILE REFERENCE: 3053-4071US2
CURRENT APPLICATION NUMBER: US/09/875,338
CURRENT FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: 60/272,107
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/209,811
PRIOR FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 698
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-875-338-9

Query Match 36.6%; Score 1249; DB 9; Length 698;
Best Local Similarity 43.8%; Pred. No. 7.3e-80;
Matches 292; Conservative 56; Mismatches 121; Indels 198; Gaps 20;

QY 32 KKGDYVELTCTASQ-KKSIQFMKNSNQIKILGNQGSFLTQKPSKLNDRADSRRLMDQ 90
Db 156 RPDYTYITICSSYGYPEAEVFMQDQGVPLGN-----VTTSQMANEQ 200
QY 91 NFPL-IINKUKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLQOGSITLTLESPP- 147
Db 201 LFDVHSILRVLVGANGTYSQ-----LVARNVLOQDAH---SSVTITPORSPTG 245
QY 148 -----GSSPVQCR-SP-----RGKN--- 162
Db 246 AVEVQPEDEPVVALVGTDAITLRCFSFSPGPSLAQLNLWQLTDYKQLVHSFTBGRQGS 305
QY 163 -----IOGKTLVSQLELQDSGTWCTTV-LONQKVEFKIDIIVLAFQKAS 208
Db 306 AYANRTALFPDILAQGNASIRLQKRVARVADBSFTCFPSIRDFGSAVSLQVAPYSKPSM 365
QY 209 SIYVKEE---GEQVESFPPLAFVYEKLTG--SGELMQAERASSSKSWITFDLKNKEVS 263
Db 366 TLBNKDLRPGDTV-----TITCSSYRGYPEAEVFMQD----- 398
QY 264 KRVYQDPKLGKMKLPLHLPLPQALPQYASSGNLTLEAKTGKLGHOENVLVYMRATQLQ 323
Db 399 -----GGQVPL-----TGAVTTSQMANEQGLF-DVHSVLRVVLGAN 433
QY 324 KNLTEVWGPSTPLMLSLKLENKAKVSKREKRVWVLANPEAGMWQCLLSDSGVLLSEN 363

Db 434 GTFYSLCLVRNP-----VLQDA-----HGSVITITQ 458
QY 384 IYVLPWTSTVPDPCPAPDEPKSCDKTHTC-----PELLGSPVLFPPKPKDTLMISTPEV 438
Db 459 PWTFP-----PEFEPEKSCDKTHTCPCPAPDELGGSPVLFPPKPKDTLMISTPEV 510
QY 439 TCVVVDVSHEDPEVKNMVVDGVEVHNAKTKRREGQYNSTYVSVLTYLHODWLNKGEY 498
Db 511 TCVVVDVSHEDPEVKNMVVDGVEVHNAKTKRREGQYNSTYVSVLTYLHODWLNKGEY 570
QY 499 KCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 558
Db 571 KCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 630
QY 559 EWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSSVMHEALHNHYTOK 618
Db 631 EWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSSVMHEALHNHYTOK 690
QY 619 SLSPG 625
Db 691 SLSPG 697

RESULT 46
US-10-077-023-9
Sequence 9, Application US/10077023
Publication No. US20030031675A1
GENERAL INFORMATION:
APPLICANT: MIKESSELL, GLEN E.
APPLICANT: CHANG, HAN
APPLICANT: FINGER, JOSHUA N.
APPLICANT: YANG, GUCHEN
APPLICANT: LU, PIN
APPLICANT: ZHOU, XIA-DI
APPLICANT: PEACH, ROBERT
TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
FILE REFERENCE: 3053-4071US3
CURRENT APPLICATION NUMBER: US/10/077,023
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 60/272,107
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: 60/209,811
PRIOR FILING DATE: 2000-06-06
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 698
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-077-023-9

Query Match 36.6%; Score 1249; DB 14; Length 698;
Best Local Similarity 43.8%; Pred. No. 7.3e-80;
Matches 292; Conservative 56; Mismatches 121; Indels 198; Gaps 20;

QY 32 KKGDYVELTCTASQ-KKSIQFMKNSNQIKILGNQGSFLTQKPSKLNDRADSRRLMDQ 90
Db 156 RPDYTYITICSSYGYPEAEVFMQDQGVPLGN-----VTTSQMANEQ 200
QY 91 NFPL-IINKUKIEDSDTYICEVEDQKEEVQLVFGLTANSDTHLQOGSITLTLESPP- 147
Db 201 LFDVHSILRVLVGANGTYSQ-----LVARNVLOQDAH---SSVTITPORSPTG 245
QY 148 -----GSSPVQCR-SP-----RGKN--- 162
Db 246 AVEVQPEDEPVVALVGTDAITLRCFSFSPGPSLAQLNLWQLTDYKQLVHSFTBGRQGS 305
QY 163 -----IOGKTLVSQLELQDSGTWCTTV-LONQKVEFKIDIIVLAFQKAS 208

Db 306 AYANRALLPDLAOGNALSRLQRYVADSGSFTCFVSIRDFGSAANSLQVAAPSKSPM 365
 Qy 209 SIIVKKE---GEQEFSPFLAFTVEKLTG--SGELMWOABASSKSWITFDLKNKEVS 263
 Db 366 TLENNKDLRPDQTV-----TTTCSYRGYPBAEFVWD----- 398
 Qy 264 KRVTQDPKLQMGKLLPLHLTLPOALPOYAGSGNLTLLAEAKTGKLDHENVLVNMRATOLQ 323
 Db 399 -----GQGVPL-----TGNTTSGMANEGQLF-DVHSLVAVLVGAN 433
 Qy 324 KNLTCGEWGPSTSPKLMLSKLENKEAKYSKREKPVWLVNPPAGMOCILSDSGVLLFSN 383
 Db 434 GTVACLVRNP-----VLQODA-----HGSVTLTGQ 458
 Qy 384 IKVLPTWSTPVPAPAPBPKCDKTHTC-----PELLGSPSVFLFPKPKDTLMSRPREV 438
 Db 459 PMTPP-----PEPEKSCDKHTHCPCPAPBELLGSPSVFLFPKPKDTLMSRPREV 510
 Qy 439 TCVVVDVSHEDPEVKFNNYVGDVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEY 498
 Db 511 TCVVVDVSHEDPEVKFNNYVGDVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEY 570
 Qy 499 KCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTLNQVSLTCLVKGFYPSDIAV 558
 Db 571 KCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTLNQVSLTCLVKGFYPSDIAV 630
 Qy 559 EWESNGQPENNYKTTTPPLDSDGSFFLYSKLTVDKSRMQGNVSCSMHEALNHYTQK 618
 Db 631 EWESNGQPENNYKTTTPPLDSDGSFFLYSKLTVDKSRMQGNVSCSMHEALNHYTQK 690
 Qy 619 SLSLSPG 625
 Db 691 SLSLSPG 697

RESULT 47

US-09-773-877A-14
 ; Sequence 14, Application US/09773877A
 ; Publication NO. US20030017977A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xia, Yu-Ping et al.
 ; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY SKIN DISEASES
 ; FILE REFERENCE: REG 710b
 ; CURRENT APPLICATION NUMBER: US/09/773,877A
 ; CURRENT FILING DATE: 2001-01-31
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 14
 ; LENGTH: 557
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Flt1(1-3 deltab)-Fc (Mut1)
 US-09-773-877A-14

Query Match 36.5%; Score 1245.5; DB 12; Length 557;
 Best Local Similarity 47.2%; Pred. No. 9.6e-80;
 Matches 295; Conservative 54; Mismatches 121; Indels 155; Gaps 24;
 Qy 85 SLMDGNPLIILIKLIEDSDTYICEVEDQKEVQLLVFGLTANSDTHLQ--GOSLTTL 143
 Db 3 SYMDTGVLCLLISCLLITGSSSGSKLDPE-----LSLKGTHIMQAGTTLH-- 51
 Qy 144 ESPGSSPSVQCR-----SPRGKNIQGGK---TLVSQLE 175
 Db 52 -----QCRGAHAKMSLPKEMVSKESRLSITKACGRN---GKQFCSTLTLLTAQ 98
 Qy 176 LODSGTWCTVTL-----QNOKVEPKIDI-----VTLAQKASSIYKKEGQVESF 223
 Db 99 ANHGFYSCKYLAVFTSKKTESAIYIFISDTGRPFPEWMSLEIPIIIMTEGR--ELVI 156
 Qy 224 PLAFVEKLT-----GSGELMWOABASSKSWITFDLKNKEVSIVRVATQD 269

Db 157 PCRVTSNITVTLKKPFLDPLIDGKRIW-----DSRKGFIIISNATYKEIGL----- 204
 Qy 270 PKLQMGKKLLHLTLPLPOALPOYAGSGNLTLLAEAKTGKLDHENVLVN-----MRATOL 322
 Db 205 -----LTCEATV-----NGHL-----YKINYLTHRQNTIIDIQISPRPVKL 242
 Qy 323 QKN-----LTCEWGPSTSPKLMLS-----LLENKEAK-----VSKREKPVWLVNPEA 365
 Db 243 LRHGTVLNLCATATPLATRRVQMTWSYPDELQDSHANIFYSVLTIDKMQ-----NKDK 296
 Qy 366 GMMOCLLSDSGQVLLBSNIKVLPWSTPVPAPAPBPKCDKTHTC-----PELLGSPSVF 420
 Db 297 GLYTCRVYR-SGSPSKSVNTSV-HIYDKAGP---GEPEKSCDKHTHCPCPAPBELLGSPSVF 351
 Qy 421 LFPKPKDTLMSITPSTPVCVVVDVSHEDPEVKFNNYVGDVEVHNAKTRPREQYNSTYR 480
 Db 352 LFPKPKDTLMSITPSTPVCVVVDVSHEDPEVKFNNYVGDVEVHNAKTRPREQYNSTYR 411
 Qy 481 VSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTLN 540
 Db 412 VSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTLN 471
 Qy 541 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPPLDSDGSFFLYSKLTVDKSRMQGN 600
 Db 472 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPPLDSDGSFFLYSKLTVDKSRMQGN 531
 Qy 601 VFSGSVNHEALNHYTQKSLSLSPG 625
 Db 532 VFSGSVNHEALNHYTQKSLSLSPG 556

RESULT 48

US-10-207-655-240
 ; Sequence 240, Application US/10207655
 ; Publication NO. US20030118592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ledbetter, Jeffrey A.
 ; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
 ; FILE REFERENCE: 390069.401C1
 ; CURRENT APPLICATION NUMBER: US/10/207,655
 ; CURRENT FILING DATE: 2002-07-25
 ; NUMBER OF SEQ ID NOS: 426
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 240
 ; LENGTH: 500
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: fusion polypeptide
 US-10-207-655-240

Query Match 36.4%; Score 1243; DB 14; Length 500;
 Best Local Similarity 47.8%; Pred. No. 1.2e-79;
 Matches 295; Conservative 37; Mismatches 127; Indels 158; Gaps 18;
 Qy 23 TQGNKVLGKGGDTVELICTASQKSIQTFWKNKSNQIKIIGNQSFLTKGPKSLNDRADS 82
 Db 27 SQSPAILASBPGEKVTWTCRASSVS--YHMYQQK-----GSSPKPIVAPSNLASGVPA 81
 Qy 83 RRLMDQG-NPPLIKLKIIEDSDTYICEVEDQKEVQLLVFGLTANSDTHLQGSGLT 141
 Db 82 RFSGSGSTYSILTSRVEADATYYCQ-----QMSFNPPTF---GAGTKL 125
 Qy 142 TLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVTLQNOKVEPKIDIIV 201
 Db 126 ELKQGGSG-----GGSGGGGS-----SQAVLQSGA-----ELV- 156
 Qy 202 LAFOKASSIYKKEGQVESFPLAFTVEKLTG--SGELMWOABASSKSWITFDLK 257
 Db 157 -----RPGASVKKMSC-----RASGTTFTSYMMHWKQTPROGLEWIG---- 193
 Qy 258 NKEVSIVRVATQDPLQMGKLLPLHLTLPOALPOYAGSGNLTLLAEAKTGKLDHENVLVN 317

```

Db 194 -----ATTPNGDPTSYNQK-RK-----211
QY 318 RATOLOKNU/TCVEWGPTSPKMLSLKENKAVERKREKRVWVWLNPEAGMOCLSDSQ 377
QY 212 -----ATLTVDKSSSTAYMOJSLTSE-----DSAVYFC-----AR 242
Db 378 VLLESN-----IKVLPWTSTPVPCEPAPEPKSCDKHTHC-----PELLGGSVFLFPKPKD 428
QY 243 VVYYSNSWYFDMWGIGTIVTVSSDOEPKSCDKHTHCPCPAPPELLGGSVFLFPKPKD 302
Db 429 TLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNATKPREEQYNSTYRVSVLTLY 488
QY 303 TLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNATKPREEQYNSTYRVSVLTLY 362
Db 489 HODMNGEKYCKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 548
QY 363 HODMNGEKYCKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 422
Db 549 KGFYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSVNH 608
QY 423 KGFYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSVNH 482
Db 609 EALHNHYTOKSLSLSPG 625
QY 483 EALHNHYTOKSLSLSPG 499
Db

RESULT 49
US-10-207-655-398
; Sequence 398, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069, 401C1
; CURRENT APPLICATION NUMBER: US/10/207, 655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 398
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-398

```

```

Query Match 36.4%; Score 1243; DB 14; Length 500;
Best Local Similarity 47.8%; Pred. No. 1,2e-79;
Matches 295; Conservative 37; Mismatches 127; Indels 158; Gaps 18;

QY 23 TQGNKVVIGKGGPVELTCTASQKSIQFMKNSNQIKLGNQGSFLTQKPSKLNDRADS 82
Db 27 SQSPALISAPGEKVTMTCASSSVS-YMHWYQKP-----GSSPKPIYAPSNLASGVRA 81
QY 83 RSLMDQG-NFPIIKNLKIEDSDTYICEVEDQKEVQLLVFGLTANSDTHLLQGSGLT 141
Db 82 RFGSGSGGTSLTISRVEAEDATATYCC-----QMSFNPPTF---GAGTKL 125
QY 142 TLSPSPSSPVQCRSRGNIGQKLSVQLQLDSDGTWCTCVLONQKVEFKIDIV 201
Db 126 ELDDGGSG-----GGSGSGGGS---SQAYLQSGA-----ELV- 156
QY 202 LAFQKASSIYKKEGEVSEFPLAFTVEKLTG-----SGELMMQAEARASSKSMITFDLK 257
Db 157 -----RPGASVKNKSC-----KASGTTFTSYMMHWKQTPROGLEMG----- 193
QY 258 NKEVSVKRVTDQPKLQMGKLLPLHLTLPOALPOYAGSGNLTLALNAKTKLHOEVNLVVM 317
Db 194 -----ATYPGNGDPTSYNQK-RK-----211

```

```

QY 318 RATOLOKNU/TCVEWGPTSPKMLSLKENKAVERKREKRVWVWLNPEAGMOCLSDSQ 377
Db 212 -----ATLTVDKSSSTAYMOJSLTSE-----DSAVYFC-----AR 242
QY 378 VLLESN-----IKVLPWTSTPVPCEPAPEPKSCDKHTHC-----PELLGGSVFLFPKPKD 428
Db 243 VVYYSNSWYFDMWGIGTIVTVSSDOEPKSCDKHTHCPCPAPPELLGGSVFLFPKPKD 302
QY 429 TLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNATKPREEQYNSTYRVSVLTLY 488
Db 303 TLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNATKPREEQYNSTYRVSVLTLY 362
QY 489 HODMNGEKYCKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 548
Db 363 HODMNGEKYCKCKVSNKALPAPIEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 422
QY 549 KGFYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSVNH 608
Db 423 KGFYPSDIAVEMESNQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRMOQGNVFCSVNH 482
QY 609 EALHNHYTOKSLSLSPG 625
Db 483 EALHNHYTOKSLSLSPG 499

```

```

RESULT 50
US-10-683-255-2
; Sequence 2, Application US/10683255
; Publication No. US20040063910A1
; GENERAL INFORMATION:
; APPLICANT: Kavanaugh, William M.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
; TITLE OF INVENTION: RECEPTOR-IMMUNOGLOBULIN FUSION
; FILE REFERENCE: PP01474, 101
; CURRENT APPLICATION NUMBER: US/10/683, 255
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 09/499, 846
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: 60/119, 002
; PRIOR FILING DATE: 1999-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-683-255-2

```

```

Query Match 36.4%; Score 1243; DB 12; Length 622;
Best Local Similarity 48.3%; Pred. No. 1,7e-79;
Matches 299; Conservative 47; Mismatches 119; Indels 154; Gaps 23;

QY 109 CEVEDQKEVQLLVFGL-TANSDTHLLQGSGLTTLSPSPSSPVQCR-RSPRGK----- 161
Db 55 CILRDVQGSINMLRQCVQLAESNRRTITGEVEVQ-DVYPADSGLACTTSPSSGSDTTY 113
QY 162 ---NIQGGKTLVSQLELDSGT-----WTCTVLONQKVEFKIDIV 202
Db 114 FSVNVSADALPSSSEDDDDSSSEKETDNTKPNPAPYWT-----SPKMEKKLHAV-- 166
QY 203 AQQKASSIYKKEGEVSEFPLAFTVE-KLQSGGLMMQAEARASSKSMITFDLKN-KE 260
Db 167 -----PAAKTIVKFKCPSSG-----TPNPTLRM-----LNQKGE 194
QY 261 VSVKRVTDQPKLQMG-----KLLPLHLTLPOALPOYAGSGNLTLALNAKTKLHOEVNL-V 315
Db 195 FK-----PDHRIGYKRYATWISIMDSVFP---SDKNYTCIIVENEGSINHHTQDLY 245
QY 316 VNRATQ---LQKNL-----TCEVWGPTSPKMLSLKLE----- 345
Db 246 VERSPHRPILQGLPANKTVLQSGNVFMCVKYSDQPHIOMLKHTEVNGSKIGPDLNLY 305

```

[illegible]

RESULT 51

```

US-10-232-838-18
; Sequence 18, Application US/10232838
; Publication No. US20030064053A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Shengliang
; APPLICANT: Martin, Jean-Francois
; APPLICANT: Liu, Dayou
; TITLE OF INVENTION: MULTIVALENT PROTEIN CONJUGATE WITH MULTIPLE LIGAND-BINDING DOMAIN
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: 26050-707
; CURRENT APPLICATION NUMBER: US/10/232,838
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US 06/316,718
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 949
; TYPE: PR1
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: MVP-C
US-10-232-838-18

```

Query Match	36.4%	Score 1243;	DB 12;	Length 949;
Best Local Similarity	38.1%	Pred. No. 3e-79;		
Matches	327;	Conservative	66;	Mismatches 150; Indels 316; Gaps 31.

Qy	33	KGDIIVELTCTASQ--KSIQFHWNKSNQIKILNGQSFLLTGPSTKLNDRADRSRLMDQG	90
Db	140	RGHTLVLTNCTATTPPLNTRVQMTWSY-----PQKNRASVRRI--DQS	181
Qy	91	N-----FPLIKLKIEDSDITYCEVEDQKEVQLVFGHTANSPHLLOGSLTTL	143
Db	182	NSHANIFYSVLTIDMQMKDKGLYTRCRSRGS-----PKSVTSHIYDKAFITVXH	234
Qy	144	ESPPGSSPVQCRSPRGKNIQGKTL---SVSQLELDGSGTWGTTVL-----	187
Db	235	GAGGGGSGG---GGSGGGGVEGAMDLIINLINSPLVSDAETSITCTACGMRPHEPITI	291
Qy	188	-----QNKUKEFKIDIVLFAFOASSIVYKKEGEQVFEPLAFVTEKLTGS---	235
Db	292	GRDFEALMNQNDPLEVTQDV---TRBAKKVYVWIRE-----KSKINGAIFC	336
Qy	236	-GELTMQA-----ERASSKSWITFDL--KNKSVYKRV---TQD-----	269
Db	337	EGRVAGEGELIRIRTMKRRQOASFPLPTLLMTYADKGDNVMSFKKVIKEEDAVIYKNGSFI	396

OY	270	-----PKLQMGKXRLPHLTLEQALPOVAG-----	SGNLTLL-----	LEAK-----	3040
OY	397	HSVPRHSEVPDILEVH--LPHAPQODAGVSARYIGGNLFTSAFTRLLYRCEAOKMGPEC			454
OY	305	-----TGKLDHVEVNLVM-----	RATOLO-----		3230
OY	455	NHLCTACMANNVGCHEDTGCICRPPGMGRTECKACALHTFRTCTCKRCSGEGCKSVFC			514
OY	324	-----KNLTCE-----	VMGP-----		333
OY	515	LPPDYGCSCATGKWKGLQCNBEACHPGFYGBDCKLRCSNNNGEMCDRFOGCLCSPGMOGLQC			574
OY	334	-----TSPKLM--LSLKLE-----	-----NKEATVSRKREKPVW-----		355
OY	575	EREGRIPMTKTVLDLDPDHLLEVNSGKFNPICKASGMPLPTNEEMTLVXKPDGTVLHPKOPNH			634
OY	360	-----VLNPEAGMMQCLLSD--SGOVLLLESNI--	KVLPTSTVPKPCAP----		392
OY	635	TDRFSVALFTIRHLPLPDSGVAVCSVNYVAGVWEKPFNISTKVLPK-----	PLNAPVID		689
OY	400	-----EPNSCDKTHTC-----	PELLGSPVFLFPKPK		428
OY	690	TGHNFAVINISSEPYFGDGPDKSKKLVDNBSKCDTHTCPCPCAPPELLGSGSVFLFPKPK			749
OY	427	KDTLMTLSRPELVYCVVVDVSHEDPEVKFMVYVDGVEVNAATKPREQVNSTYRVSVLT			488
OY	750	KDTLMTLSRPELVYCVVVDVSHEDPEVKFMVYVDGVEVNAATKPREQVNSTYRVSVLT			809
OY	487	VLDHODMNGKEYCKVSNKALPAPLEKTIISAKAQPREQVYTLPPSRDELTKNQVSLTC			548
OY	810	VLDHODMNGKEYCKVSNKALPAPLEKTIISAKAQPREQVYTLPPSRDELTKNQVSLTC			865
OY	547	LVKGFYPSDIAVEMESNGQPENNYKTTPPVLDSGSPFLYSKLTVDXSRWQGNVFCSV			606
OY	870	LVKGFYPSDIAVEMESNGQPENNYKTTPPVLDSGSPFLYSKLTVDXSRWQGNVFCSV			922
OY	607	MHEALNNHTQKSLSLSPG	625		
OY	930	MHEALNNHTQKSLSLSPG	948		

RESULT 52

```

US-09-875-338-5
; Sequence 5: Application US/09875338
; Patent No. US20020095024A1
; GENERAL INFORMATION:
; APPLICANT: MIKESELL, GLEN E.
; APPLICANT: CHANG, HAN
; APPLICANT: FINGER, JOSHUA N.
; APPLICANT: YANG, GUCHEN
; APPLICANT: LU, PIN
; APPLICANT: ZHOU, XIA-DI
; APPLICANT: PEACH, ROBERT
; TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
; TITLE OF INVENTION: IMMUNOMODULATION
; FILE REFERENCE: 3053-4071US2
; CURRENT APPLICATION NUMBER: US/09/875.338
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/272,107
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/209,811
; PRIOR FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-875-338-5

```

Query Match 36.4%; Score 1242.5; DB 9; Length 480;
 Best Local Similarity 59.2%; Pred. No. 1.3e-79;
 Matches 276; Conservative 33; Mismatches 70; Indels 87; Gaps 17;

```

219 VEFSPFLAFTVEKLTGSGEL--WMOAERASSSKSWTF-----DLKNKEVSVK---RYVQ 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
42 IECKFP---VEKQDLALIVWEME---DKNIIFQHGEBDLKVQSHSYQRARLAK 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
269 DPKLQMGKKLPHLT-----LPQALPOYAGSG--NLTLALEATGKLIHQEVLVNR 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
94 D-QLSIG-NALQITVQLQDAGVYRCMISYGADYKRIIVKVNAPYKINQRI-LVDP 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
319 ATQLQKNTLCEVWG-PTSPKMLSLKENKAV---SKREKPV----- 359
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 VTS-EHELTQAEQYPAEYVITWSSDHQVLSGKTTTNSKREKLFNVTSLRLINTTNE 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
360 -----VINPAGMOCCLSDSGVLLSENIKVLPTWSTPVPSP-----APEPKSC 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 IFYCTFRRLDPE-----ENHTAEVLVPEPLAHFPRNERTRGDPEEPKSC 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405 DKHTTC----PELGGPSVFLFPPPKKDTLMSRTPEVTCVVVDVSHEDPEVKFMWYD 459
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
254 DKHTTCPPCPAPPELLGGPSVFLFPPPKKDTLMSRTPEVTCVVVDVSHEDPEVKFMWYD 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
460 GVEVHNAKTPREEQYNSTYRVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTSKAK 519
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 GVEVHNAKTPREEQYNSTYRVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTSKAK 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
520 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 579
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
374 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 DGSFPLYSKLTVDKSRMOCGNVFSCSYMEHALHNHYTQKSLSLSPG 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 434 DGSFPLYSKLTVDKSRMOCGNVFSCSYMEHALHNHYTQKSLSLSPG 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 53
 US-10-077-023-5
 ; Sequence 5, Application US/10077023
 ; Publication No. US20030031675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MIKESILL, GLEN E.
 ; APPLICANT: CHANG, HAN
 ; APPLICANT: FINGER, JOSHUA N.
 ; APPLICANT: YANG, GUCHEN
 ; APPLICANT: LIU, PIN
 ; APPLICANT: ZHOU, XIA-DI
 ; APPLICANT: PEACH, ROBERT
 ; TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
 ; FILE REFERENCE: 3053-4071US3
 ; CURRENT APPLICATION NUMBER: US/10/077,023
 ; CURRENT FILING DATE: 2002-02-15
 ; PRIOR APPLICATION NUMBER: 60/272,107
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/209,811
 ; PRIOR FILING DATE: 2000-06-06
 ; NUMBER OF SEQ ID NOS: 138
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 480
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: fusion construct
 US-10-077-023-5

Query Match 36.4%; Score 1242.5; DB 14; Length 480;
 Best Local Similarity 59.2%; Pred. No. 1.3e-79;
 Matches 276; Conservative 33; Mismatches 70; Indels 87; Gaps 17;

```

219 VEFSPFLAFTVEKLTGSGEL--WMOAERASSSKSWTF-----DLKNKEVSVK---RYVQ 268
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
42 IECKFP---VEKQDLALIVWEME---DKNIIFQHGEBDLKVQSHSYQRARLAK 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
269 DPKLQMGKKLPHLT-----LPQALPOYAGSG--NLTLALEATGKLIHQEVLVNR 318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
94 D-QLSIG-NALQITVQLQDAGVYRCMISYGADYKRIIVKVNAPYKINQRI-LVDP 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
319 ATQLQKNTLCEVWG-PTSPKMLSLKENKAV---SKREKPV----- 359
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
151 VTS-EHELTQAEQYPAEYVITWSSDHQVLSGKTTTNSKREKLFNVTSLRLINTTNE 209
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
360 -----VINPAGMOCCLSDSGVLLSENIKVLPTWSTPVPSP-----APEPKSC 404
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
210 IFYCTFRRLDPE-----ENHTAEVLVPEPLAHFPRNERTRGDPEEPKSC 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
405 DKHTTC----PELGGPSVFLFPPPKKDTLMSRTPEVTCVVVDVSHEDPEVKFMWYD 459
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
254 DKHTTCPPCPAPPELLGGPSVFLFPPPKKDTLMSRTPEVTCVVVDVSHEDPEVKFMWYD 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
460 GVEVHNAKTPREEQYNSTYRVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTSKAK 519
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 GVEVHNAKTPREEQYNSTYRVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTSKAK 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
520 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 579
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
374 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTPPVLD 433
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 DGSFPLYSKLTVDKSRMOCGNVFSCSYMEHALHNHYTQKSLSLSPG 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 434 DGSFPLYSKLTVDKSRMOCGNVFSCSYMEHALHNHYTQKSLSLSPG 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 54
 US-10-207-655-15
 ; Sequence 15, Application US/10207655
 ; Publication No. US20030118592A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ledbetter, Jeffrey A.
 ; APPLICANT: Hayden-Ledbetter, Martha S.
 ; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
 ; FILE REFERENCE: 390069, 401C1
 ; CURRENT APPLICATION NUMBER: US/10/207,655
 ; CURRENT FILING DATE: 2002-07-25
 ; NUMBER OF SEQ ID NOS: 426
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 15
 ; LENGTH: 499
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: MOUSE-HUMAN HYBRID FUSION PROTEIN
 ; NAME/KEY: SITE
 ; LOCATION: (1) .. (265)
 ; OTHER INFORMATION: MOUSE ANTI-HUMAN CD20 SCFV: 2H7
 ; NAME/KEY: DOMAIN
 ; LOCATION: (266) .. (499)
 ; OTHER INFORMATION: HUMAN IGG1 WILD TYPE HINGE, CH2, CH3 FC
 US-10-207-655-15

Query Match 36.4%; Score 1242.5; DB 14; Length 499;
 Best Local Similarity 47.5%; Pred. No. 1.3e-79;
 Matches 294; Conservative 33; Mismatches 129; Indels 163; Gaps 18;

```

23 TQGNKVLGKKDPTVETCTASQKSIQPHMKNNSNIGIKLNGGSLTGKPSGLNLRADS 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
27 SOSPALISPGSKVMTTCRASSSVS-YWMMYQOKP-----GSPKRWIYAPSLASGVPA 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
83 RSLMDQG-NPFLIINKLIEPSDYICEVEDQKEEVQLLVFGLTANSDTHLQGGSLTL 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 82 RPSGSGSTSYSLTISRVEADATYYCQ-----QMSFNPPTF---GAGTKL 125
Qy 142 TLSPSPGSPVQCRSPRGNKNIQGGKTLVSQLELDQSGTWTCTVLQKQKVEFKIDIV 201
Db 126 ELKGGGSG-----GGSGGGGS---SQAYLQSGA-----ELV- 156
Qy 202 LAFKASSIYKKEGEQVEFSPLATFVEKLTG---SGELMWOAERASSSKSWITPDLK 257
Db 157 -----RPGASVKNMSC-----KASGYTFTSYNMHWKQTPROGLEWIG----- 193
Qy 258 NKEVSVKRVTPDPLQNGKKLPLHLTLPLQALPOYAGSGNLTLEAKTGKLGHEVNLVVM 317
Db 194 -----ALPVGNDTSTYNOKF-GK----- 211
Qy 318 RATOLQKRLTCGEWGPSTPKLMSLKLKNEK---AKYSKREKRVWVLNPPAGMQLLS 373
Db 212 -----ATLTVDKSSSTAYMQLSLTSEDNAVYFCARVYVSNSTYVFD----- 254
Qy 374 DSGVLLSNNIKVLPTW--STPVPCEPAPRPSCKDTHTC-----PELLGSPSVLFPPKP 426
Db 255 -----VWGTTTVVSDQEPKSCDKTHCPPCAPPELLGSPSVLFPPKP 299
Qy 427 KDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNAAKTRPEQYNSTYRVSVLT 486
Db 300 KDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNAAKTRPEQYNSTYRVSVLT 359
Qy 487 VLHODWLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTLC 546
Db 360 VLHODWLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTLC 419
Qy 547 LVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDSKRWQOGNVFSCSV 606
Db 420 LVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDSKRWQOGNVFSCSV 479
Qy 607 MHEALHNHYTQKSLSLSPG 625
Db 480 MHEALHNHYTQKSLSLSPG 498

RESULT 55
US-10-207-655-148
Sequence 148, Application US/10207655
Publication No. US20030118592A1
GENERAL INFORMATION:
APPLICANT: Ledbetter, Jeffrey A.
TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
FILE REFERENCE: 390069.401C1
CURRENT APPLICATION NUMBER: US/10/207,655
CURRENT FILING DATE: 2002-07-25
NUMBER OF SEQ ID NOS: 426
SOFTWARE: PatentIn version 3.0
SEQ ID NO 148
LENGTH: 499
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Mouse-Human hybrid fusion protein
US-10-207-655-148

Query Match 36.4%; Score 1242.5; DB 14; Length 499;
Best Local Similarity 47.5%; Pred. No. 1.3e-79;
Matches 294; Conservative 33; Mismatches 129; Indels 163; Gaps 18;

Qy 23 TQGNKVVLGKKGDVLTCTASQKSIQFHWKNSNOIKLGNQGSFLTGPSKLNDRADS 82
Db 27 SQSPALISAPGKVTMTCTASSSVS-YMHWYQKP-----GSSPKMTIYAPSNLASGVA 81
Qy 83 RRLSLMDQG-NPPLIIKLIKIDSDTYICEVEDQKEVQVLVFGLTANSQTHLQGSGLT 141
Db 82 RPSGSGSTSYSLTISRVEADATYYCQ-----QMSFNPPTF---GAGTKL 125
Qy 142 TLSPSPGSPVQCRSPRGNKNIQGGKTLVSQLELDQSGTWTCTVLQKQKVEFKIDIV 201

Db 126 ELKGGGSG-----GGSGGGGS---SQAYLQSGA-----ELV- 156
Qy 202 LAFKASSIYKKEGEQVEFSPLATFVEKLTG---SGELMWOAERASSSKSWITPDLK 257
Db 157 -----RPGASVKNMSC-----KASGYTFTSYNMHWKQTPROGLEWIG----- 193
Qy 258 NKEVSVKRVTPDPLQNGKKLPLHLTLPLQALPOYAGSGNLTLEAKTGKLGHEVNLVVM 317
Db 194 -----ALPVGNDTSTYNOKF-GK----- 211
Qy 318 RATOLQKRLTCGEWGPSTPKLMSLKLKNEK---AKYSKREKRVWVLNPPAGMQLLS 373
Db 212 -----ATLTVDKSSSTAYMQLSLTSEDNAVYFCARVYVSNSTYVFD----- 254
Qy 374 DSGVLLSNNIKVLPTW--STPVPCEPAPRPSCKDTHTC-----PELLGSPSVLFPPKP 426
Db 255 -----VWGTTTVVSDQEPKSCDKTHCPPCAPPELLGSPSVLFPPKP 299
Qy 427 KDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNAAKTRPEQYNSTYRVSVLT 486
Db 300 KDTLMISRTPEVTCVVVDVSHEDPEVKFNNYVDSGEVHNAAKTRPEQYNSTYRVSVLT 359
Qy 487 VLHODWLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTLC 546
Db 360 VLHODWLNGEKYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTLC 419
Qy 547 LVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDSKRWQOGNVFSCSV 606
Db 420 LVKGFYPSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDSKRWQOGNVFSCSV 479
Qy 607 MHEALHNHYTQKSLSLSPG 625
Db 480 MHEALHNHYTQKSLSLSPG 498

RESULT 56
US-10-053-530-15
Sequence 15, Application US/10053530
Publication No. US2003013939A1
GENERAL INFORMATION:
APPLICANT: Ledbetter, Jeffrey
TITLE OF INVENTION: Binding Domain-Immunoglobulin Fusion Proteins
FILE REFERENCE: 390069.401
CURRENT APPLICATION NUMBER: US/10/053,530
CURRENT FILING DATE: 2002-01-17
PRIORITY FILING DATE: 2001-01-17
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn version 3.0
SEQ ID NO 15
LENGTH: 499
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: MOUSE-HUMAN HYBRID FUSION PROTEIN
NAME/KEY: SITE
LOCATION: (1)..(265)
OTHER INFORMATION: MOUSE ANTI-HUMAN CD20 SCFV: 2H7
NAME/KEY: DOMAIN
LOCATION: (266)..(499)
OTHER INFORMATION: HUMAN IGG1 WILD TYPE HINGE, CH2, CH3 FC
US-10-053-530-15

Query Match 36.4%; Score 1242.5; DB 14; Length 499;
Best Local Similarity 47.5%; Pred. No. 1.3e-79;
Matches 294; Conservative 33; Mismatches 129; Indels 163; Gaps 18;

Qy 23 TQGNKVVLGKKGDVLTCTASQKSIQFHWKNSNOIKLGNQGSFLTGPSKLNDRADS 82
Db 27 SQSPALISAPGKVTMTCTASSSVS-YMHWYQKP-----GSSPKMTIYAPSNLASGVA 81


```

Query Match          36.4%; Score 1242; DB 16; length 634;
Best Local Similarity 45.5%; Seed No. 2e-79;
Matches 298; Conservative 39; Mismatches 122; Indels 196; Gaps 18;

QY      10 LLLVQLAL-LP-----AATGNGKVKLGKGDVELTCTASQKSIOFTM---K 54
       |||||         |||||         |||||         |||||         |||||
Db      2  LLVTSLLLCELPHPAFLIPDIQTITSSLSASLIGRVITSCRASDISRYLWVGQK 61
       |||||         |||||         |||||         |||||         |||||

QY      55 NSNQIKILGNQSFLTGG-PSKLNDRADSRKSLMDQGNFPLIKVLKIEDSTTYCEVED 113
       ::|||         ::|||         ::|||         ::|||         ::|||

```

```

      RESULT 58
US-09-875-338-17
; Sequence 17, Application US/09875338
; Patent No. US20020095024A1
; GENERAL INFORMATION:
; APPLICANT: MIKESSELL, GLEN E.
; APPLICANT: CHANG, HAN
; APPLICANT: FINGER, JOSHUA N.
; APPLICANT: YANG, GUOCHEN
; APPLICANT: LU, PIN
; APPLICANT: ZHOU, XIA-DI
; APPLICANT: PEACH, ROBERT
; TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
; FILE OF INVENTION: IMMUNOMODULATION
; FILE REFERENCE: 3053-4071US2
; CURRENT APPLICATION NUMBER: US/09/875,338
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/872,107
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/709,811
; PRIOR FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 17
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: fusion construct
US-09-875-338-17
Query Match          36.4%; Score 1241.5; DB 9; Length 451;
Best Local Similarity 50.0%; Pfd. No. 1.4e-79;
```



```

Query Match      36.3%; Score 1239.5; DB 12; Length 555;
Best Local Similarity 45.1%; Pred. No. 2.5e-79;
Matches 296; Conservative 55; Mismatches 167; Indels 139; Gaps 18

QY      4 GVEPRHLLVLTALALPAATQG-----NKKVLGKKDPTVELTCTAOKKSIO 50
Dp      2 GAPAAASLLLLLLLPACCMAPGAAVLSDDSQDQPMWTSDEFVVA--GCTVVLKCOVDHEDSS 59
QY      51 FHWKNSQIKIL-----GNQGSFLTKGSKLNDRAADSRSLMDGNFLLIKNLKIE 102
Dp      60 LQWNPAAQOQLTYFGEKRALRDNRLOLVTSPTPHELS-----ISISNVALA 103
QY      103 DSDPTVIE-----VEDQKEEYQILVF---GLTASDTHLQGGOSLTLTLESPPGSSPFAVQ 154
Dp      104 DEGGYTSIFPMPTAKSLVTVTIGTIOKPLITTKYKSLRKQDTATLNCQS--SGSKPAPAR 162
QY      155 CRSPRGKNIQGGKLTSVSOLEQDSGTWCTVLQNKQKVEPKIDIVLAFQKASIVYKK 214
Dp      163 LTKRKGQDELHGEPTRIQ---EDPNKGTFTV-----SSSVTRQV 198
QY      215 EGEQVESSFLPAFVEKLTGSGELMWAERASSSKSMITPDLKNKEVSVKRYTQD--PKLQ 273
Dp      199 TREDDGASIVCSVNHESLKG-----ADRTSORIEVLVY-----TPTAMIRPDPPHPR 245
QY      274 MGKTLPHLTLPLQALPQVAGSGNLTALAEAKTGHQELNVLMWRATOLQKNTICEVNGP 333
Dp      246 EGGKLLHLC-----EGRGN-----PVPOQLMEKSGS 272
QY      334 TSPKLMLSLKLENKAVSKREKEPVWVNLNPEAGMOCLLSDSGOVLDESNTIKVLEPTWSP 393
Dp      273 VPP-----LKMTOBSALIFP-----PINKSDSGYTGCTATSN-----MGSYKAYVTLTAVN 317
QY      394 VPCGAPERPKSCDKTKTC-----PELLGPAVSFLPPPKQDTLMTSRPEVVCVVVDVSH 448
Dp      318 DPSVPSPGSGSKDKHTHTPCPCAPALLGQPSVFLPPPKQDTLMTSRPEVVCVVVDVSH 377
QY      449 DPEVKFMVYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQMDLNGKEKYCKVSNKALP 508
Dp      378 DPEVKFMVYDGVGVHNAKTKPREEQYNSTYRVVSVLTVLHQMDLNGKEKYCKVSNKALP 437
QY      509 APIEKTISKAKGQREPOVYTLTPSRDELTKNOVSLTCLVKGYFSPSDIAVEMESNGQDEN 568
Dp      438 VPIEKTISKAKGQREPOVYTLTPSRREMTKNQVSLTCLVKGYFSPSDIAVEMESNGQDEN 497
QY      569 NYKTTPTVLDSDGSFFLYSKLTVDKSRMQQGNVSCSMYHEALHNHYTQKSLSLSPG 625
Dp      498 NYKTTPTVLDSDGSFFLYSKLTVDKSRMQQGNVSCSMYHEALHNHYTQKSLSLSPG 554

RESULT 62
US-10-120-198B-2
; Sequence 2, Application US/101020198B
; Publication No. US20030215427A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Michael
; TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
; FILE REFERENCE: 1954-337
; CURRENT APPLICATION NUMBER: US/10/120,198B
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/282,859
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 631
; TYPE: PRT
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: mouse-human chimera
US-10-120-198B-2

```

Best Local Similarity 45.6%; Pred. No. 3e-79;
Matches 295; Conservative 47; Mismatches 116; Indels 189; Gaps 20;

```
OY 5 VPFRLLVLQALILPAATQGNKIVLGGKGDVELTCTASQKSIQF--HM---KNSQI 59
DB 13 LHPAFLLIPQVQLOQPGAE---LVKPGASVKLSKSCASGTTGYMHHVKKQPGGL 67
OY 60 KILNQSFLTKPSKLNDRADSRRLW-DQGNFP--LIINKLKIEDDTYICEVEDQKE 116
DB 68 EMIGINP--SNGRRNVERFRSKATLVDSSTAFQGLSTEDNAVFCARD----- 121
OY 117 EVQLVPELTIANSDFHLIQGSLITLESPPSSPSVQCRSPRGKNIQGGKTLVSQ--- 173
DB 122 ---YVGTSYNFD---YMGQDTLTVSSGGGSGG---GSGGGGSDIQWTQSSS 166
OY 174 ---LELDQSGTWTCTVLONOKKVEPKIDIVILAPQKASSIYKKEGEVEFSPLATVE 230
DB 167 SFSVSLGDRVITTC-----KANEID1----- 186
OY 231 KLTVSGELMWAERASSSKSWITFDLKNKEVSVKKRTQDPKLQMGKLLPLHLTPQALPQ 290
DB 187 ---NNRLAWQQTPTGNS-----PRLISGATNLVTGVS---R 218
OY 291 YAGSNTLTALBAKTKLHQEVNLVWBRATOLQKNLTCEVWGPTSPKMLSLKLENKAK 350
DB 219 FSGSG-----SGKDY---TLTITSLQAE----- 238
OY 351 VSKREKPVWLNPEKGMQCLSDSGVLLSNNIKVLPWTSTPVPCPA-----PEPKS 403
DB 239 ---DPAITYC-----QQWSTPFTGSGTELEIKVEKKS 269
OY 404 CDKHTTC-----PELLGSPVFLPPPKPDITLMSRTEPVTCVVVDVSHEDPEVKFNYY 458
DB 270 SDKHTTCPCAPPELLGSPVFLPPPKPDITLMSRTEPVTCVVVDVSHEDPEVKFNYY 329
OY 459 DGEVHNAKTKPREQVNSTYRVVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTISKA 518
DB 330 DGEVHNAKTKPREQVNSTYRVVSVLTVLHQMNLNGEKYCKVSNKALPAPIEKTISKA 389
OY 519 KQGREPQVYTLPPSRDELTKQVSLTCLVKGFPSPDAVWESNGQENNYKTTTPVLD 578
DB 390 KQGREPQVYTLPPSRDELTKQVSLTCLVKGFPSPDAVWESNGQENNYKTTTPVLD 449
OY 579 SDGSFFLYSKLTVDSKRWQGNVFSQVMHEALNHYTKSLSPG 625
DB 450 SDGSFFLYSKLTVDSKRWQGNVFSQVMHEALNHYTKSLSPG 496
```

RESULT 63

US-10-363-427-24
; Sequence 24, Application US/10363427
; Publication No. US20030195338A1
; GENERAL INFORMATION:
; APPLICANT: MedexGen Inc.
; APPLICANT: CHUNG, Yong Hoon
; APPLICANT: HAN, Ji Woong
; APPLICANT: LEE, Hye Ja
; APPLICANT: CHOI, Eun Yong
; APPLICANT: KIM, Jin Mi
; APPLICANT: YIM, Soo Bin
; TITLE OF INVENTION: Concatametric Immunoadhesion
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/363,427
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Kapatentin 1.71
; SEQ ID NO 24
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-363-427-24

Query Match 36.3%; Score 1239; DB 14; Length 502;

Best Local Similarity 47.9%; Pred. No. 2.4e-79;
Matches 293; Conservative 40; Mismatches 111; Indels 168; Gaps 18;

```
OY 22 ATQGNKIVLGGKGDVELTCTASQKSIQFHMKNKSNQIKIINGQSFLTKPSKLNDRAD 81
DB 50 ATEVRVTVLRAQADSQVEVCAATY-----MMGNELTF-----LDDISIC 87
OY 82 SRSLLMDQGNFPILLINKLKIEDSDTYICEVEDQKEVQLVFGLTANSDFHLIQGSLTL 141
DB 88 TGTSSGNQVN--LTIQGLRAMDTGLYICKVE----- 116
OY 142 TLESPPSSPSVQCRSPRGKNIQGGKTLVSQLELDQSGTWTCTVLONOKKVEFKIDIV 201
DB 117 -LWVPP-----PYLIDIGNGTQGVNDTE-----PCNDSDDNNHTAQ---PAVV 155
OY 202 LAFQKA-SIIYKKEGEVEFSPLATVEKLTSGELMWAERASSSKSWITFDLKNKE 260
DB 156 LASSRGIASFV-----CEYASPGKATEVAT---VLRAQADSQVEVCAATY----- 198
OY 261 VSVKRVITQDPKLQMGKLLPLHLTPQALPQYAGSNTLTALBAKTKLHQEVNLVWBRAT 320
DB 199 ---MMGNELTF---LDDISICTGSSGN-----QVNLTIQGLR 229
OY 321 QLOKNL--TCEVWGPTSPKMLSLKLENKAKVSKREKPVWVNLNPEAGMOCCLSDSGQVL 379
DB 230 AMDTGVLICVKELMWPPRYLGI-----GNGTQIYVIDPE----- 264
OY 380 LESNIKVLPTWSTVPCP-APEPKSCDKHTTC-----PELLGSPVFLPPPKPDITLMS 433
DB 265 ---PCPDSAPPKSCDKHTTCPCAPPELLGSPVFLPPPKPDITLMS 309
OY 434 RTEPVTCVVVDVSHEDPEVKFNMYVDGEVHNAKTKPREQVNSTYRVVSVLTVLHQMNL 493
DB 310 RTEPVTCVVVDVSHEDPEVKFNMYVDGEVHNAKTKPREQVNSTYRVVSVLTVLHQMNL 369
OY 494 NGKEYKCKVSNKALPAPIEKTISKAKQPREPQVYTLPPSRDELTKQVSLTCLVKGFP 553
DB 370 NGKEYKCKVSNKALPAPIEKTISKAKQPREPQVYTLPPSRDELTKQVSLTCLVKGFP 429
OY 554 SDIAVWESNGQENNYKTTTPVLDSDGSFFLYSKLTVDSKRWQGNVFSQVMHEALHN 613
DB 430 SDIAVWESNGQENNYKTTTPVLDSDGSFFLYSKLTVDSKRWQGNVFSQVMHEALHN 489
OY 614 HYTKSLSPG 625
DB 490 HYTKSLSPG 501
```

RESULT 64

US-10-404-724-8
; Sequence 8, Application US/10404724
; Publication No. US20030203447A1
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Arnold H.
; TITLE OF INVENTION: Methods and Materials for Increasing Expression of Recombinant
; FILE REFERENCE: 13698US01
; CURRENT APPLICATION NUMBER: US/10/404,724
; PRIOR APPLICATION NUMBER: US 60/368,530
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 8
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-404-724-8

Query Match 36.3%; Score 1238.5; DB 12; Length 465;
Best Local Similarity 46.1%; Pred. No. 2.4e-79;
Matches 296; Conservative 34; Mismatches 101; Indels 211; Gaps 17;

[illegible]

```

; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/733,764
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/590,656
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-357-653-2
```

```

Query Match 36.2%; Score 1237; DB 14; Length 704;
Best Local Similarity 81.4%; Pred. No. 5.3e-79;
Matches 240; Conservative 9; Mismatches 12; Indels 34; Gaps 5;
```

```

Qy 360 VLNPAGMOCCLSD-SCGVLLSNI--KVLPTWSTPVPCCAP----- 399
Db 414 ILPDSGVWCSVNTVAGVKEKFNISVKLPK-----PLANPNVIDTGHNFVAINISSE 468
Qy 400 ----EPKSCDKTHTC-----PELLGSPVFLPPPKPKDTLMI SRTPEVTCVVVDVSHEDP 450
Db 469 PYFGEPKSCDKTHTCPCPAPPELLGSPVFLPPPKPKDTLMI SRTPEVTCVVVDVSHEDP 528
Qy 451 EVKFNMYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHODMLNGKEYCKCKVSNKALPAP 510
Db 529 EVKFNMYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHODMLNGKEYCKCKVSNKALPAP 588
Qy 511 IEKTSKAKGPREPOVYTLPPSRDELTKQVSLTCLVKGFPSPDI AVEWESNGQPENNY 570
Db 589 IEKTSKAKGPREPOVYTLPPSRDELTKQVSLTCLVKGFPSPDI AVEWESNGQPENNY 648
Qy 571 KTTPEVLDSDGSFPLYSKLTVDKSRMOQGNVFCSCVWHEALHNHYTKSLSPG 625
Db 649 KTTPEVLDSDGSFPLYSKLTVDKSRMOQGNVFCSCVWHEALHNHYTKSLSPG 703
```

RESULT 67

```

US-10-282-162-52
; Sequence 52, Application US/10282162
; Publication No. US20030143697A1
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
; FILE REFERENCE: AND USING
; CURRENT APPLICATION NUMBER: US/10/282,162
; PRIOR FILING DATE: 2002-10-28
; PRIOR APPLICATION NUMBER: 09/787,835
; PRIOR FILING DATE: 1999-09-22
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 52
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-282-162-52
```

```

Query Match 36.2%; Score 1237; DB 14; Length 915;
Best Local Similarity 44.7%; Pred. No. 7.5e-79;
Matches 293; Conservative 51; Mismatches 129; Indels 182; Gaps 18;

Qy 33 KGDVVELTCT-----ASOKKSIOFHW-KNSNOIKILGNQSFILTKGPSKLNDRADSR 84
Db 380 EGEPAALRCQVPRWMLASVPRINLTMHKDSARTVG-----BEET 422
Qy 85 SLMDQGNPLIILKLIKIEDSDTYICEVED-----QKEVQVLLVFGITANS DTHL---LOGQ 137
Db 423 RMMWADGALMLLPLAQ-EDSGTYVCTTNNASYCDKMSIELRVF---ENTDAFLPFIISYPQ 478
```

```

Qy 138 SLTLT-----LESPPGS-----SPVQCRSPRGKNIQGGKTLVSQLEQ 177
Db 479 ILTITSGVLYCPPLSETRTKTDVKIOWMYDSILLDNDNKFSLVSRGTHLVHVDVLE 538
Qy 178 DSGTWCTV-----LONQKVEFKIDIVLAFOKASIVYKKEGOVEF 221
Db 539 DAGYRCVLTFAHGGQOYNITRSIELRIKKKEETIPVILISPLTISASLSR----- 591
```

```

Qy 222 SFPLATVEKLTGSGE-----LMMQAEASSSKSMITFDLKNKSVSVRYVQDPKLOMGK 276
Db 592 ---LTIPCKVFLGTGTPLTMLMWTANDTHIESAV-----PGSRVTEGPPROEYSE 638
Qy 277 KLPLHLTPQALPOYAGSGNLTALAEATGK-LHOEVNLVYMRATQLOKNLTCEVWGPTS 335
Db 639 NNENVIEVP-----LIFDPYTRDLMDFRCVHNLTSPQ----- 673
Qy 336 PKMLSLKENKEAKVSRKREKPVVNLPEAGMOCCLSDSGVLLSNIKVLPTWSTPVP 395
Db 674 ---TLRTVKEAS----- 683
Qy 396 CPAPEPKSCDKTHTC-----PELLGSPVFLPPPKPKDTLMI SRTPEVTCVVVDVSHEDP 450
Db 684 ---STSGDKTHTCPCPAPPELLGSPVFLPPPKPKDTLMI SRTPEVTCVVVDVSHEDP 739
Qy 451 EVKFNMYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHODMLNGKEYCKCKVSNKALPAP 510
Db 740 EVKFNMYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHODMLNGKEYCKCKVSNKALPAP 799
Qy 511 IEKTSKAKGPREPOVYTLPPSRDELTKQVSLTCLVKGFPSPDI AVEWESNGQPENNY 570
Db 800 IEKTSKAKGPREPOVYTLPPSRDELTKQVSLTCLVKGFPSPDI AVEWESNGQPENNY 859
Qy 571 KTTPEVLDSDGSFPLYSKLTVDKSRMOQGNVFCSCVWHEALHNHYTKSLSPG 625
Db 860 KTTPEVLDSDGSFPLYSKLTVDKSRMOQGNVFCSCVWHEALHNHYTKSLSPG 914
```

RESULT 68

```

US-10-679-620-62
; Sequence 62, Application US/10679620
; Publication No. US20040110930A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology
; APPLICANT: Reini, Stephen J
; TITLE OF INVENTION: MULTIMERIC PROTEIN ENGINEERING
; FILE REFERENCE: 34150-004A
; CURRENT APPLICATION NUMBER: US/10/679,620
; PRIOR FILING DATE: 2003-10-03
; PRIOR APPLICATION NUMBER: 60/415,940
; PRIOR FILING DATE: 2002-10-03
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent In Ver. 3.2
; SEQ ID NO 62
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: p9E10chimericv1-1, see Example 15
; US-10-679-620-62
```

```

Query Match 36.2%; Score 1236; DB 16; Length 715;
Best Local Similarity 46.4%; Pred. No. 6.3e-79;
Matches 285; Conservative 39; Mismatches 104; Indels 186; Gaps 15;

Qy 30 LGGKDVVELTCTAS-----OKKSIOFHWKNSNOIKILGNQSFILTKGPSKL 76
Db 269 LVKGGSLKLSKASGTFPSHYGMSWVRQTPDKRLN-----VATISRGTY-THYDSV 322
Qy 77 NDRADSRSLMDQGNPLIILKLIKIEDSDTYICEVEDQKEVQVLLVFGITANS DTHLLOQ 136
Db 323 KGRFTISRDN-DKNALYLOMNSLSKSEDTAMYC-----ARRSEFYVYGNTYYVSAMDYWG 376
```

```

QY 137 QSLTLTLESPSSSPVQCRSPRGKNIQGGKTLTSSVQLDSDGCTWTCTVLOKQKVERK 196
DB 377 QGASVTSSASTKGPSPVFPPLAPSSKSTSGG-TNALCCL----- 413
QY 197 IDIVLAFQKASSIVVKKKEGQVEFSFPLAFTVEKLTGSGELMWQARASSSKMTTFDL 256
DB 414 -----VKDYFPEPVTVS-----MNSGALTSG----- 434
QY 257 KKEKSVKRVTPDKLQMGKLPFLHLTLPOALPQVYSGNLTLALAKTGKLGHOENVLV 316
DB 435 -----VH-TFPAVL-QSSGLTSSSVTVVPSSTLGTQYI-- 467
QY 317 MRATQLOKNTLCEWGPSTSPKMLSLKLEKKAQVSKREKPYVVLNPEAGMOCCLSDSG 376
DB 468 -----CNV-----NHKPSNTKVDKRV----- 483
QY 377 QVLLSNIKVLPTWSTPVPCEPAPPEKSCDKTHTC-----PELLGSPVFLFPKPKDTLM 431
DB 484 -----EPKSCDKTHTCPCEPAPPELLGSPVFLFPKPKDTLM 520
QY 432 ISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHOD 491
DB 521 ISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYRVSVLTVLHOD 580
QY 492 WLNKGEYKCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKQVSLTCLVKGF 551
DB 581 WLNKGEYKCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKQVSLTCLVKGF 640
QY 552 YPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGNVSCSVMEAL 611
DB 641 YPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGNVSCSVMEAL 700
QY 612 HNNHYTKSLSLSPG 625
DB 701 HNNHYTKSLSLSPG 714

```

```

RESULT 69
US-10-683-255-6
; Sequence 6, Application US/10683255
; Publication No. US20040063910A1
; GENERAL INFORMATION:
; APPLICANT: Kavanaugh, William M.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
; FILE REFERENCE: PP01474.101
; CURRENT APPLICATION NUMBER: US/10/683,255
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 09/499,846
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: 60/119,002
; PRIOR FILING DATE: 1999-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-683-255-6

```

```

Query Match 36.2%; Score 1235.5; DB 12; Length 497;
Best Local Similarity 66.2%; Pred. No. 4.2e-79;
Matches 255; Conservative 18; Mismatches 54; Indels 58; Gaps 8;
QY 293 GSGNLTALAEAKTGKLIHQ---EVLNVVRATOLQKNL--TCEVWGPTSPKMLSLKLE-- 345
DB 118 GSIHNTYQLDVERSHPRILOGLPANKTVALGSNVEPMCKVSDQPHIOMLKHEVN 177
QY 346 -----NKEAKVSKREKPYVVLN-----PEAGMOCCLSDSGQVLLSNIK 385
DB 178 GSKIGPDLNLPYVQILKTAGVNTTDEKMEVHLRNVSFEDAGEYTCLAGNSIGLSHHS--- 234

```

```

QY 386 VLPTWST-----VPCPAP-----EPKSCDKTHTC-----PELLGSPVFL 420
DB 235 ---AMLTVLEALEERPAWMTSPLYLEGSSPGLQEPKSCDKTHTCPCEPAPPELLGSPVFL 291
QY 421 LFPKPKDTLMSRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYR 480
DB 292 LFPKPKDTLMSRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYR 351
QY 481 VVSVTLVHODWLNGEKYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKN 540
DB 352 VVSVTLVHODWLNGEKYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKN 411
QY 541 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGN 600
DB 412 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGN 471
QY 601 VFSCSVMEHALHNNHYTKSLSLSPG 625
DB 472 VFSCSVMEHALHNNHYTKSLSLSPG 496

```

```

RESULT 70
US-10-683-255-4
; Sequence 4, Application US/10683255
; Publication No. US20040063910A1
; GENERAL INFORMATION:
; APPLICANT: Kavanaugh, William M.
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
; FILE REFERENCE: PP01474.101
; CURRENT APPLICATION NUMBER: US/10/683,255
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 09/499,846
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: 60/119,002
; PRIOR FILING DATE: 1999-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-683-255-4

```

```

Query Match 36.2%; Score 1235.5; DB 12; Length 525;
Best Local Similarity 66.2%; Pred. No. 4.5e-79;
Matches 255; Conservative 18; Mismatches 54; Indels 58; Gaps 8;
QY 293 GSGNLTALAEAKTGKLIHQ---EVLNVVRATOLQKNL--TCEVWGPTSPKMLSLKLE-- 345
DB 146 GSIHNTYQLDVERSHPRILOGLPANKTVALGSNVEPMCKVSDQPHIOMLKHEVN 205
QY 346 -----NKEAKVSKREKPYVVLN-----PEAGMOCCLSDSGQVLLSNIK 385
DB 206 GSKIGPDLNLPYVQILKTAGVNTTDEKMEVHLRNVSFEDAGEYTCLAGNSIGLSHHS--- 262
QY 386 VLPTWST-----VPCPAP-----EPKSCDKTHTC-----PELLGSPVFL 420
DB 263 ---AMLTVLEALEERPAWMTSPLYLEGSSPGLQEPKSCDKTHTCPCEPAPPELLGSPVFL 319
QY 421 LFPKPKDTLMSRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYR 480
DB 320 LFPKPKDTLMSRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREEOYNSTYR 379
QY 481 VVSVTLVHODWLNGEKYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKN 540
DB 380 VVSVTLVHODWLNGEKYCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKN 439
QY 541 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGN 600
DB 440 QVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPFLYSKLTVDKSRMOQGN 499

```


Qy 601 VFSCSVMEALHNHYTOKSLSPG 625
Db 500 VFSCSVMEALHNHYTOKSLSPG 524

RESULT 71
US-10-207-655-397
; Sequence 397, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 397
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-397

Query Match 36.1%; Score 1234; DB 14; Length 500;
Best Local Similarity 47.6%; Pred. No. 5.4e-79;
Matches 294; Conservative 37; Mismatches 128; Indels 158; Gaps 18;
Qy 23 TQGNKVVVGGKQDPTVELTCTASQKSIQFMKNSNQIKLGNQSGFLTKGSKLNDRAD 82
Db 27 SQSRAILSASFGKVTMTCRASSVS-YMHVYQKP-----GSSRPMTYASNLASGPA 81
Qy 83 RRSIMDQG-NPFLIIKNLIKEDSDTYICEVEDQKEVQLLVFGLTANSDTLLQGQSLTL 141
Db 82 RFSGSGSGTSLTISRVEAEDAATVYCO-----QMSFNPPTF---GAGTKL 125
Qy 142 TLSPRSGSPVQCRSPRGKNIQGGKITLSVQLELDGSGTCTVYLQKQKVERKIDIV 201
Db 126 ELKQGGGSG-----GGSGGGGS---SQAYLQSGA-----ELV- 156
Qy 202 LAFQASSIVYKKEGEQVFSFPLAFVTEKLTG-----SGELMQAERASSSKMTFPLK 257
Db 157 -----REGASVYKSC-----KASGTFITSYNMHVYKQTPRQGLEWIG----- 193
Qy 258 NKEVSVKRVTDPKLQMGKLLPLHLTPQALPOYAGSGNLTALAEAKTGKLEHVNILVVM 317
Db 194 -----AIVPQNGDTSYNQKFK-GK----- 211
Qy 318 RATOLQKLTCEWNGPFSPKMLSLKLENKEAKVSKREKPVWVLNPEAGMQCLSDSGQ 377
Db 212 -----ATLTVDKSSSTAYMQLSLTSE-----DSAVVFC-----AR 242
Qy 378 VLLESN-----IKVLPTMSTPVPCCAPPEKSCDKTHT-----CPELLGGSPVFLFPKPKD 428
Db 243 VVYISNSNMYRPDWGCTGTTTVVSSDQPKSCDKHTHTSPCCAPPELLGGSPVFLFPKPKD 302
Qy 429 TLMISRPPEVTCVVVDVSHEDPEVKFMNYVVGVEVHNAKTRPREQVNSTYRVVSVLTVL 488
Db 303 TLMISRPPEVTCVVVDVSHEDPEVKFMNYVVGVEVHNAKTRPREQVNSTYRVVSVLTVL 362
Qy 489 HQDWLNGEKYCKYKSNKALPAPIETKISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 548
Db 363 HQDWLNGEKYCKYKSNKALPAPIETKISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLV 422
Qy 549 KGFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFLLSKLTIVKSRMQQGNVSCSYMH 608
Db 423 KGFPSDIAVEMESNGQPENNYKTTTPVLDSDGSFLLSKLTIVKSRMQQGNVSCSYMH 482
Qy 609 EALHNHYTOKSLSPG 625
Db 483 EALHNHYTOKSLSPG 499

RESULT 72
US-10-452-646-9
; Sequence 9, Application US/10452646
; Publication No. US20040018593A1
; GENERAL INFORMATION:
; APPLICANT: Carton, Jill M.
; APPLICANT: Steadnet, Kimberly C.
; APPLICANT: Scallion, Bernard J.
; APPLICANT: Jili, Giles-Romar
; TITLE OF INVENTION: ANTI-RELP FUSION ANTIBODIES, COMPOSITIONS, METHODS AND USES
; FILE REFERENCE: CEN0296 NP
; CURRENT APPLICATION NUMBER: US/10/452,646
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: US 60/385,305
; PRIOR FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 367
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-452-646-9

Query Match 36.1%; Score 1233; DB 15; Length 367;
Best Local Similarity 74.7%; Pred. No. 4.2e-79;
Matches 251; Conservative 17; Mismatches 30; Indels 38; Gaps 10;
Qy 326 LTCVWGPSTP-KLMLSLKLENKEAKV---SKREPVWV--LNPE-AGMQCL----- 371
Db 33 LECQSYNGAHLASILSLKEASTIAEYISGYRSQPIWIGLHDPQKQOQWIDGAMVLY 92
Qy 372 LSDSGVLT-----LESNIKVLPTWSTPVPCC-----APPKSCDKTHTC----- 410
Db 93 RSMGKMGNGKHCHEMSSNNFL-TWSSN-ECNKRQHFLLCKYRPEPKSCDKTHTCPPCP 150
Qy 411 -PELLGSPVFLFPKPKDITLMSRTPREVTCVVVDVSHEDPEVKFMNYVVGVEVHNAKTK 469
Db 151 APELLGGSPVFLFPKPKDITLMSRTPREVTCVVVDVSHEDPEVKFMNYVVGVEVHNAKTK 210
Qy 470 PREQVNSTYRVVSVLTVLHQDWLNGEKYCKYKSNKALPAPIETKISKAKQPREPQVYTL 529
Db 211 PREQVNSTYRVVSVLTVLHQDWLNGEKYCKYKSNKALPAPIETKISKAKQPREPQVYTL 270
Qy 530 LPPSRDELTKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPVLDSDGSFLLSKL 589
Db 271 LPPSRDELTKNQVSLTCLVKGFPYSDIAVEMESNGQPENNYKTTTPVLDSDGSFLLSKL 330
Qy 590 TVDKSRMQQGNVSCSYMHLEALHNHYTOKSLSPG 625
Db 331 TVDKSRMQQGNVSCSYMHLEALHNHYTOKSLSPG 366

RESULT 73
US-10-207-655-396
; Sequence 396, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 396
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion polypeptide

US-10-207-655-396

```

Query Match      36.1%; Score 1233; DB 14; Length 500;
Best Local Similarity 47.6%; Pred. No. 6,4e-79;
Matches 294; Conservative 37; Mismatches 128; Indels 158; Gaps 18;

QY 23 TOGNKVLGKGGTVELTCTASQKSIQFHWKNSNQIKLGNQGSFLTKGPKSLANDRAD 82
DB 27 SOSPAILASPEKERTWTCRASSSVS-YMHMYQKP---GSPKRWIYAFSLASGVPA 81
QY 83 RSLMDQ--NFPILIKNLKIEDSDTYICEVEDOKEEVOLVFGITANSPTHLQGSSTL 141
DB 82 RFGSGSGSTYSLTLSVEAEDATYYCQ-----QMSNPPTF---GAGTGL 125
QY 142 TLESPPGSSPVOCSEPRGNIOGKTLISVQLFLDSDGTWTCVLOKQKKEFKDIYV 201
DB 126 ELKDGGGSG-----GGSGGGGS---SQAYLQSGA-----ELV- 156
QY 202 LAFQKASSIYKKEGQVEFSPLAFVTEKLTG---SGELMWQBARASSSKSWITFDLK 257
DB 157 -----RPGASVKNKSC-----KASGTTTISYMHVYKQTPROGLEWIG--- 193
QY 258 NKEVSVKRYTQDPKLOMGKPLHLTLPCALPOYAGSGNLTALAEKTKLHOEVNLVYM 317
DB 194 -----AIPGNGDTSYNGKFK-GK----- 211
QY 318 RATQLOKNTLCEVWGPSTPKLMLSLKENKAVSKREKPVVVLNPEAGMOCLLSDSQ 377
DB 212 -----ATLVKSSSTAYWQLSSLTSE-----DSAVYFC-----AR 242
QY 378 VLLESN---IKVLPMTSTVPVPCAPPEKSCDKTHTC-----PELLGSPVFLPPPKPD 428
DB 243 VVYSSSYMYFVDWGTGTTTVSSDDEPKSSDKTHTCPCPAPELLGSPVFLPPPKPD 302
QY 429 TLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEOYNSTYRVVSVLTVL 488
DB 303 TLMISRTPEVTCVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEOYNSTYRVVSVLTVL 362
QY 489 HQMLNGKEYKCYVSNKALPAPIEKTSKAKGQPREQVYTLPPSRDELTKNOVSLTCLV 548
DB 363 HQMLNGKEYKCYVSNKALPAPIEKTSKAKGQPREQVYTLPPSRDELTKNOVSLTCLV 422
QY 549 KGYPSDIAVEMESNGOPENNYKTPVLPDSDSFFLYSKLTVDKSRMOQGNVFCSSVMH 608
DB 423 KGYPSDIAVEMESNGOPENNYKTPVLPDSDSFFLYSKLTVDKSRMOQGNVFCSSVMH 482
QY 609 EALHNYTKSLSPG 625
DB 483 EALHNYTKSLSPG 499

RESULT 74
US-10-363-427-20
; Sequence 20, Application US/10363427
; Publication No. US20030195338A1
; GENERAL INFORMATION:
; APPLICANT: MedexGen Inc.
; APPLICANT: CHONG, Yong Hoon
; APPLICANT: HAN, Ji Woong
; APPLICANT: LEE, Hye Ja
; APPLICANT: CHOI, Eun Yong
; APPLICANT: KIM, Jin Mi
; APPLICANT: YIM, Soo Bin
; TITLE OF INVENTION: Concatametric Immunoadhesion
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/363,427
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Kopatentin 1.71
; SEQ ID NO 20
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-363-427-20

```

Query Match      36.1%; Score 1233; DB 14; Length 502;
Best Local Similarity 68.5%; Pred. No. 6,4e-79;
Matches 250; Conservative 12; Mismatches 43; Indels 60; Gaps 7;

QY 290 QVAGSGNLTALAEATGKLGHOEVNLVYMRATQLOKNTLCTC-----EYWGPT 334
DB 168 EYASFOKAT-----EYRYTVLRQADSVTEVCATAYYMGNELTFLDSDICTGT 215
QY 335 SPKMLSLKENKA-----KVSKEKPVVNLNPEAGMOCCLSDSGOVLLESNIRY 386
DB 216 SSGNQVNLITQGLRAMDTGLYICKVLEWPPRYLIGIGNG-----TQIYV 260
QY 387 LPTWSTVPCP-APPEKSCDKTHTC-----PELLGSPVFLPPPKPDITMISRTPEVTC 440
DB 261 I-----DPEPCPDASBPSCDKTHTCPCPAPELLGSPVFLPPPKPDITMISRTPEVTC 316
QY 441 VVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEOYNSTYRVVSVLTVLHQDMLNGKEYKC 500
DB 317 VVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEOYNSTYRVVSVLTVLCHQDMLNGKEYKC 376
QY 501 KVSNNKALPAPIEKTSKAKGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEM 560
DB 377 KVSNNKALPAPIEKTSKAKGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEM 436
QY 561 ESNQOPENNYKTPVLPDSDSFFLYSKLTVDKSRMOQGNVFCSSVMHALLHNYTKSL 620
DB 437 ESNQOPENNYKTPVLPDSDSFFLYSKLTVDKSRMOQGNVFCSSVMHALLHNYTKSL 496
QY 621 SLSPG 625
DB 497 SLSPG 501

RESULT 75
US-10-207-655-346
; Sequence 346, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069, 401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 346
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-346

Query Match      36.1%; Score 1233; DB 14; Length 543;
Best Local Similarity 46.0%; Pred. No. 7,1e-79;
Matches 296; Conservative 43; Mismatches 115; Indels 190; Gaps 19;

QY 1 NMRGVFRRLLVLQALLPATQGNKVVLGKGGTVELTCTASQKSIQFHWKNSNQIK 60
DB 19 MSRGVD-----IVL-----TQSPITTAABGENTITTCRASSSVSYMYWQKS--- 62
QY 61 ILGNQGSFLTKGPKSLANDRADSRSLMDQ--NFPILIKNLKIEDSDTYICEVEDOKEEVQ 119
DB 63 --GAPKMLTYDPTSKLASGVPMRFGSGSGTYSIALINMETRBDATYYCQ----- 111
QY 120 LAVFGITANSPTHLQGSSTLTLBSPGSSSPVOCSEPRGNIOGKTLISVQLFLDSD 179
DB 112 -----QMSSTPLTF-----GSGTLEIRKGGGGGGGGGGGGGGQVOLDKEA 152
QY 180 G-----TWTCVTVLOKQKVEFKIDIVLAFQKASSIYVKKGEQVEFSPLAFV 229

```

```
Db 153 GPGIYVPTQTSLTCTV-----SGFS----- 173
Qy 230 EKLNGSELWMQARASSSSKSWITFDLKNKEVSKRVYTOPDKLQMGKPLPLHLTPQALP 289
Db 174 --LTSVGVHW-----IRQP---GKLEW---MGII 196
Qy 290 QYASGNLTLLAEKTKGLHGEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKEA 349
Db 197 YVDCGTVDYNSAIKSR-----LSTSRDTS-----KSGVFLKINSIQ- 231
Qy 350 KVSREKREKRVWVNLNEAGWMQCL---LSDSGVLLSNIKVLPTWSTVPVCPAPEKSCDK 406
Db 232 -----TDDTAMVYCARLHPDYWGQ-----GVWVTVSSDL-----EPKSSDK 267
Qy 407 THT-----CPELLGSPVFLFPPPKKDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVYDGV 461
Db 268 THTSPPSPAPPELLGSSSVFLFPPPKKDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVYDGV 327
Qy 462 EVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 521
Db 328 EVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQ 387
Qy 522 PREQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPEVLDSDG 581
Db 388 PREQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPEVLDSDG 447
Qy 582 SFFLYSKLTVDKSRWQGNVFCSCVMHEALNHTYOKSLSLSPG 625
Db 448 SFFLYSKLTVDKSRWQGNVFCSCVMHEALNHTYOKSLSLSPG 491
```

RESULT 76

```
US-10-107-991B-3
; Sequence 3, Application US/10107991B
; Publication No. US20040058445A1
; GENERAL INFORMATION:
; APPLICANT: LEDBETTER, JEFFREY
; APPLICANT: HAYDEN-LEDBETTER, MARTHA
; APPLICANT: HELLSTROM, INGBERD
; APPLICANT: HELLSTROM, KARL ERIK
; TITLE OF INVENTION: ACTIVATION OF TUMOR-REACTIVE LYMPHOCYTES VIA ANTIBODIES
; FILE REFERENCE: 034474.0004
; CURRENT APPLICATION NUMBER: US/10/107, 991B
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: 60/286, 585
; PRIOR FILING DATE: 2001-04-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 2.1
; SEQ ID NO 3
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mouse-Human Fusion Protein
; NAME/KEY: SIGNAL
; LOCATION: (1)..(23)
; OTHER INFORMATION: L6 V kappa signal peptide
; NAME/KEY: DOMAIN
; LOCATION: (24)..(133)
; OTHER INFORMATION: g19-4 mouse anti-human CD3 light chain variable domain
; NAME/KEY: PEPTIDE
; LOCATION: (134)..(148)
; OTHER INFORMATION: (Gly4Ser)3 linker peptide
; NAME/KEY: DOMAIN
; LOCATION: (149)..(270)
; OTHER INFORMATION: g19-4 mouse anti-human VH domain
; FEATURE:
```

```
; NAME/KEY: MISC FEATURE
; LOCATION: (271)..(504)
; OTHER INFORMATION: human IgG1 Fc domain (hinge, CH2, CH3)
; FEATURE:
; NAME/KEY: TRANSMEM
; LOCATION: (505)..(555)
; OTHER INFORMATION: human CD80 transmembrane domain and cytoplasmic tail
US-10-107-991B-3
```

```
Query Match 36.1%; Score 1233; DB 12; Length 555;
Best Local Similarity 46.3%; Pred. No. 7.3e-79;
Matches 299; Conservative 42; Mismatches 123; Indels 182; Gaps 20;
```

```
Qy 1 MNRGVPRHLVLVQLALPLAATQGNKVVGLKGGDTVELTCTASQKKSIOFHW---KNSN 57
Db 19 MSRGVDIQ-----MTQTSLSASLSDRTTSCRASQDRIANLWYQOQRPDG 65
Qy 58 QIKILNGSGSFLTKGPSKLNDRADSRSLWDQ--NPELLIKNLKIEDSDTYICEVEDQKE 116
Db 66 TVKLK---LYYT---SRLHSVPSRFGSGSGTDYSLTIANLPEDIATYFCQ----- 112
Qy 117 EVQLLIVGLTANSPTHLQGSQSLTTLTSPSPGSSPSYQCSPRGKNIQGGKT---LSVQ 173
Db 113 -----QENTLPWTF---GGGTKLVTKRELGGSGGGSGGSGSIDE 151
Qy 174 LEIODSGTWTCTVLQNGKVFPKIDIVLAFQKASSIVYKKEGQVFFPLAFTVEKLT 233
Db 152 VOLQSGP-----ELV---KQASMSCASSG---YSF-TGYIVV--- 183
Qy 234 GSGELMWQARASSSSKSWITFDLKNKEVSKRVYTOPDKLQMGKPLPLHLTPQALPYAG 293
Db 184 -----WLKSHGKNLEWIGLINPYKGLT---TYNKKFK----- 213
Qy 294 SGNLTLLAEKTKGLHGEVNLVVMRATOLQKNTLCEVWGPTSPKMLSLKLENKEAVSK 353
Db 214 -GKATLTVDKSSSTAYWE-----LSTSEDSAVYYCA 245
Qy 354 REK-----PYVNLNPEAGMWQCLLSDSGVLLSNIKVLPTWSTVPVCPAPEKSC 404
Db 246 RSGYDGSDMYFDFW---GAGTTVTYSSDL-----EPKSS 277
Qy 405 DKHT-----CPELLGSPVFLFPPPKKDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVD 459
Db 278 DKHTSPSPAPPELLGSSSVFLFPPPKKDTLMTSRTEVTCVVVDVSHEDPEVKFNNYVD 337
Qy 460 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 519
Db 338 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 397
Qy 520 GQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPEVLD 579
Db 398 GQPREPQVYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPEVLD 457
Qy 580 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALNHTYOKSLSLSPG 625
Db 458 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALNHTYOKSLSLSPG 503
```

RESULT 77

```
US-10-363-427-16
; Sequence 16, Application US/10363427
; Publication No. US20030195338A1
; GENERAL INFORMATION:
; APPLICANT: MedexGen Inc.
; APPLICANT: CHUNG, Yong Hoon
; APPLICANT: HAN, Ji Woong
; APPLICANT: LEE, Hye Ja
; APPLICANT: CHOI, Eun Yong
; APPLICANT: KIM, Jin Mi
; APPLICANT: YIM, Soo Bin
; TITLE OF INVENTION: Concatametric Immunoadhesion
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/363, 427
```

```

; CURRENT FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Kopacentic 1.71
; SEQ ID NO 16
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-363-427-16

Query Match      36.1%; Score 1232; DB 14; Length 377;
Best Local Similarity 68.8%; Pred. No. 5.1e-79;
Matches 251; Conservative 12; Mismatches 42; Indels 60; Gaps 7;

QY 290 QYAGSNLTALAEAKGKGLHGVNLYVMATQLQKSLT-----EVMGFL 334
DB 43 EVASPKAT-----EVRVTVLRQADSQVTEVCATYMGNEELTFLLDSICTGT 90
QY 335 SPKMLSLKLENKA-----KVSREKPVVWVLPBAGMQLLSDSQVLLSEINIKY 386
DB 91 SSGNQVNLITQGLRANDTGLYICKVELMPPYYLIGNG-----TQIVY 135
QY 387 LPTWSTPVPBP-APEPKSCDKHTC-----PELLGSPVFLFPPPKDPTLMSRTPBVT 440
DB 136 I-----DPEPCPDAAEPKSCDKHTCPCPAPBELLGSPVFLFPPKDKDTLMSRTPBVT 191
QY 441 VVVVDSHEDPEVKFNMYVDGVEVNAKTKPREQYNSTRVSVLTVLHODMLNGEKYK 500
DB 192 VVVVDSHEDPEVKFNMYVDGVEVNAKTKPREQYNSTRVSVLTVLHODMLNGEKYK 251
QY 501 KVSNNKLPAPLEKTIKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 560
DB 252 KVSNNKLPAPLEKTIKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 311
QY 561 ESNQGENNYKTPPVLDSDGSFELYSKLTVDKSRMQGNVFSQVMEALHNHYTQKSL 620
DB 312 ESNQGENNYKTPPVLDSDGSFELYSKLTVDKSRMQGNVFSQVMEALHNHYTQKSL 371
QY 621 SLSPG 625
DB 372 SLSPG 376

RESULT 78
US-10-334-235-38
; Sequence 38; Application US/10334235
; Publication No. US20040131591A1
; GENERAL INFORMATION:
; APPLICANT: Oxford Biomedica (UK) Ltd.
; APPLICANT: Kingman, Alan
; APPLICANT: Bebbington, Christopher
; APPLICANT: Carroll, Miles
; APPLICANT: Eillard, Fiona
; APPLICANT: Kingman, Susan
; APPLICANT: Myers, Kevin
; APPLICANT: Lamikandra, Abigail
; TITLE OF INVENTION: VECTOR SYSTEM
; FILE REFERENCE: 53268200920
; CURRENT APPLICATION NUMBER: US/10/334,235
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 10/060,585
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: PCT/GB00/04317
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US 09/445,375
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide of 575aa

```

```

US-10-334-235-38

Query Match      36.1%; Score 1231; DB 16; Length 600;
Best Local Similarity 44.7%; Pred. No. 1.1e-78;
Matches 302; Conservative 49; Mismatches 131; Indels 194; Gaps 21;

QY 30 LKKKGDVTELTCTAS--QKSIQFMKNSNQKIIG-----NQSFLTGPSKLANDRA-- 80
DB 33 LVKPGASVKISCKASGYSTGYMHWKSHKSLFEMIRIPNNGVTLTYNQKFKDKALL 92
QY 81 ---DSRSIMDQGNPPLIINKLIKEDSPDYCE-----VEDQKEVOLLVFG----- 124
DB 93 TVDKSSTIAY-----MELRSLSSEDSAYVCARSMITNNYMDWGQVTSVTYSSCGGG 146
QY 125 -----LRASDTALL--QGOSLTLT-----LSBPQSSPS--VQ 154
DB 147 SGGGGTGGGSSIVWTFPTFLVAGDVTITTCASQGSNDVAVYQKPGQSPFLIS 206
QY 155 CSRPR-----QKNIQGGKTLISVQLQDSGTMTCTVLAQNO-----KVEFKID 198
DB 207 YTSRIRAGVPDRFISGCTDFTTISTLQAEIDLAVFQODYNSPPTGGGKLEIK-- 264
QY 199 IVVLAFOKASSIVYKKEGQVEFSEPLAFTVEKLTSGLGLMWQAERASSKSWITFDLKN 258
DB 265 -----RAST-----KPSV--FPL-----APSKS----- 282
QY 259 KEVSVKRVYQDPKLGKGLPLHLTLPOALPOYAG---SGNLTALAEATGKLHGVNLY 315
DB 283 -----TSGGTALGCLVKGDPFEPVTVSNWNGALTSGVHTFPVAVLQSSGLYSLSVY 334
QY 316 VNRATQL-QKNLTCEVWGTPSPKMLSLKENKEKVASREKPVWVLPBAGMQLLSD 374
DB 335 TVPSSSLGTQYICNV-----NHKSNNTVQKXV----- 363
QY 375 SGQVLLSEINIKVLTPTWSTPVPBPAPPEPKSCDKHTC-----PELLGSPVFLFPPPKDPT 429
DB 364 -----EPKSCDKHTCPCPAPBELLGSPVFLFPPPKDPT 398
QY 430 LMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVNAKTKPREQYNSTRVSVLTVLH 489
DB 399 LMISRTPEVTCVVVDVSHEDPEVKFNMYVDGVEVNAKTKPREQYNSTRVSVLTVLH 458
QY 490 QDMNGEKYKCVSKKALPAPLEKTIKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 549
DB 459 QDMNGEKYKCVSKKALPAPLEKTIKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVK 518
QY 550 GFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMQGNVFSQVME 609
DB 519 GFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFELYSKLTVDKSRMQGNVFSQVME 578
QY 610 ALHNHYTQKSLSLSPG 625
DB 579 ALHNHYTQKSLSLSPG 594

RESULT 79
US-10-435-299-7
; Sequence 7; Application US/10435299
; Publication No. US20040052783A1
; GENERAL INFORMATION:
; APPLICANT: Weimer, George
; APPLICANT: Gingsrich, Roger
; APPLICANT: Link, Brian
; APPLICANT: Tso, J. Yun
; TITLE OF INVENTION: HUMANIZED ANTIBODIES AGAINST CD3
; FILE REFERENCE: 05882-0176-CNUS04
; CURRENT APPLICATION NUMBER: US/10/435,299
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 09/618,380
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 08/397,411
; PRIOR FILING DATE: 1995-03-01
; PRIOR APPLICATION NUMBER: US 07/859,583

```

```

; PRIOR FILING DATE: 1992-03-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Complete heavy chain of Humanized 1D10 Ab
US-10-435-299-7

```

```

Query Match      36.0%; Score 1230.5; DB 12; Length 446;
Best Local Similarity 46.7%; Pred. No. 8.2e-79;
Matches 284; Conservative 36; Mismatches 103; Indels 185; Gaps 14;

```

```

QY 30 LGKGDVVELTCTSAQKSIQF--HMKNSNQIKILGNQSFYLTGSPSKLNDRAISRSL- 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 11 LKPSSETLSTCTGSGFLTYGVWVRQSPKGLWIGVMGSGSTENNAFTSRLLTIS 70
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 87 --WPGNPELTIKLIKIEDSDTYICEVEDQKEVQLVFGLTANSDTHLQ--GQSLLT 142
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 71 KDTSKNQVSLKLNSTLTADTAATYTC-----AKNDRYAMDYWGQGLTLYT 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 143 LESPSSPSYVQCSPRGKNIQGGKTLVSQLELDQSGTWCTVYLQNKVEFKIDIVVL 202
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 114 VSSASTKGPSVFLPAPSSKSTSGG-TALGCL----- 144
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 203 AFQKASSIVYKKEGQVEFSPFLAFTVEKLTGSGELMWQAEKASSSKWITFDLKNKVS 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 145 -----VDYFPEPEVTVS-----WNSGALTSG----- 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 263 VKRVTQDPKLGKKLPLHLTLPLQALPOYAGSGNLTLLAEKTKLHQEVNLVYMRATQL 322
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 166 -----VH-TFPAVL-QSSGLYLSVVTVPSLSLGTQYI----- 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 323 QKNLTCEWGPSTPKMLSLKLENKAVSKREKPVWVLNPEAGMOCILSDSGVLLS 382
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 199 -----CNV-----NHKPSNTKVDKKV----- 214
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 383 NIKVLPWTSTVPCPAPPSKCDKTHTC-----PELLGSPSVFLPPPKDITLMSRTE 437
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 215 -----EPKSCDKHTHPCPAPPELLGSPSVFLPPPKDITLMSRTE 257
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 438 VTCVVDVSHDEPVEKFMWYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHQMNGKE 497
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 258 VTCVVDVSHDEPVEKFMWYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHQMNGKE 317
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 498 YKCAVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIA 557
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 318 YKCAVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIA 377
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 558 VEMESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQGNVFGSCVMHEALAHNYTQ 617
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 378 VEMESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQGNVFGSCVMHEALAHNYTQ 437
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 618 KSLSLSPG 625
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 438 KSLSLSPG 445
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 80
US-10-207-655-270
; Sequence 270, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 270
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion polypeptide
US-10-207-655-270

```

```

Query Match      36.0%; Score 1230.5; DB 14; Length 550;
Best Local Similarity 45.1%; Pred. No. 1.1e-78;
Matches 299; Conservative 35; Mismatches 108; Indels 221; Gaps 19;

```

```

QY 1 MNRGVPFRHLVLVQLALPLPAATQGNKVLGKKDVELTCTASQ--KKS-I-QFMKNSN 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 19 MSRGVDIQ-----MTQSPSSLSASLGKTTIKTSQDIIKTIQWQHPRK 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 58 QIKILGNQSFYLTG--PSKLNDRASRSLMDQNPFLIKLIKIEDSDTYICEVEDQKE 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 66 GPRLLIYVSTLQGISRFSGSGSGR-----DYSLISRLNLEPEDIAATYCO----- 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 117 EVQLVFGVLTANSDTHLQGSLLTLESPSSPSYVQCSPRGKNIQGGKTLVSQLEL 176
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 113 -----QYDNLPLTF-----GSGTKLEIRGGGSGGSGGSDVQL 150
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 177 QDSG-----TWCTVYLQNKVEFKIDIVLAFQKASSIVYKKEGQVEFSPFLA 226
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 151 QDSGFLVKSQSLSLTCV----- 170
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 227 FVEKLTG-----SGELM-WQAEKASSSKW--ITFDLKN-----KEVSVKRVTDPK 271
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 171 -----TGYSITSGFYNNWIRQFPGNKLEWNGHISHDGRNNVPSLINRISITRDT-S-K 222
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 272 LQNGKKLPLHLTLPLQALPOYAGSGNLTLLAEKTKLHQEVNLVYMRATQLQKNLT 327
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 223 NQFELKLSVTTEDTATYFCARHYSGAM----- 252
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 328 CEWVGPTSPKMLSLKLENKAVSKREKPVWVLNPEAGMOCILSDSGVLLSNIKVL 387
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 253 -DYWG----- 267
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 388 PTWSTVPCPAPPSKCDKTHTC-----CPPELLGSPSVFLPPPKDITLMSRTEVTCV 442
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 268 -----EPKSCDKHTHSPSPAPPELLGSSVFLPPPKDITLMSRTEVTCV 315
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 443 VDVSHDEPVEKFMWYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHQMNGKEKYCKV 502
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 316 VDVSHDEPVEKFMWYVDGVEVHNAAKTPREEOYNSTRVVSVLTVLHQMNGKEKYCKV 375
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 503 SNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMES 562
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 376 SNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMES 435
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 563 NGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQGNVFGSCVMHEALAHNYTQSL 622
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 436 NGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRMQGNVFGSCVMHEALAHNYTQSL 495
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 623 SPG 625
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 496 SPG 498
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 81
US-10-418-836-38
; Sequence 38, Application US/10418836
; Publication No. US20040018573A1
; GENERAL INFORMATION:
; APPLICANT: Power, Scott D.
; APPLICANT: Wang, Huaming
; APPLICANT: Ward, Michael
; TITLE OF INVENTION: Production of Functional Antibodies in Filamentous Fungi
; FILE REFERENCE: GC741-2
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 836
; SOFTWARE: PatentIn version 3.0

```

```

? CURRENT FILING DATE: 2003-04-17
? PRIOR APPLICATION NUMBER: US 60/373,889
? PRIOR FILING DATE: 2002-04-18
? PRIOR APPLICATION NUMBER: US 60/411,540
? PRIOR FILING DATE: 2002-09-18
? PRIOR APPLICATION NUMBER: US 60/452,134
? PRIOR FILING DATE: 2003-03-04
? PRIOR APPLICATION NUMBER: US 60/411,537
? PRIOR FILING DATE: 2002-09-18
? NUMBER OF SEQ ID NOS: 40
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 38
?
? LENGTH: 972
?
? TYPE: PRT
?
? ORGANISM: Artificial Sequence
?
? FEATURE:
?
? OTHER INFORMATION: fusion protein
?
? US-10-418-836-38

```

Query Match	36.0%	Score 1230.5;	DB 15;	Length 972;
Best Local Similarity	46.7%	Pred. No. 2.4e-78;		
Matches 284; Conservative	36;	Mismatches 103;	Indels 185;	Gaps 14;

0Y	30	LGKGDVTELTCTASQKSIQF--HWKSNQIKILGNQSGFLTKPCKLNDRABSRL-	86
Db	537	LVKSEETSLTCTVSGFSLTNGVMWVROSPEKGLEMIIVKWSGGSTERYNAFISRLTIS	596
0Y	87	--WQGNFPLIIFKNKIEDSPDYICEVDEQKEEVQLVLFGLTANSDTHLQ--GGSLTLT	142
Db	597	KDTSKNQVSLKLNLSLTADTAIVYC-----ARNDRTAMDYWGCGTLVT	639
0Y	143	LESPPGSSPYQCRSPRGKNIQGGKTLVSQLELODSGTWCTCYLQNKQKVEFKIDIVL	202
Db	640	VSSASTKQSPVFLPAPSSTKSTSGG--TAAIGCL-----	670
0Y	203	AFQKASSTVYKKEGQVFEFPFLATFTVEKLTGSGELMMQABRASSKSMITFDLKNKEVS	262
Db	671	-----VKDYFPEPVTVS-----WMSGALTSG-----	691
0Y	263	VKRVTDPEKLOMGKKLPLHLTLPOALPOYAGSGNLTLALCAKTLKHOBVNLVVRATOL	322
Db	692	-----VH--TFPAIL--QSSGILYSLSVTVTPSSSLGQTYI-----	724
0Y	323	QKNLTCEYWGPTSPKPLMLSLKLENKEAYKSREKPEVWVNLNPEAGMIOCLSDSGOVLLES	382
Db	725	-----CNY-----NHKPSNTKVYDKKV-----	740
0Y	383	NIKVLPTMSTPVCAPAPBPBKSCDKTHTC-----PELLGSPSYFLFPKPXKDTLMTSRPTE	437
Db	741	-----BPKSCDKTHTCPCCAPAPELGSPSYFLFPKPXKDTLMTSRPTE	783
0Y	438	VTCVVVDVSHEDPEVKFPMVYVDGVEVHNAKTKPRBEQYNSTYRVYSVLTVLIHQDLNNGKE	497
Db	784	VTCVVVDVSHEDPEVKFPMVYVDGVEVHNAKTKPRBEQYNSTYRVYSVLTVLIHQDLNNGKE	843
0Y	498	YKCKRSNKAALPAPIKTKISKAKGQREPOVYTLTPSSRDELTKNOVSLTCLYKGFPPSDIA	557
Db	844	YKCKRSNKAALPAPIKTKISKAKGQREPOVYTLTPSSRDELTKNOVSLTCLYKGFPPSDIA	903
0Y	558	VEMESNGQPENNYKTTTPVLYSDGSGFFLYSKLTVPKSRQOQGNVPSCSYMHGALNNHTQ	617
Db	904	VEMESNGQPENNYKTTTPVLYSDGSGFFLYSKLTVPKSRQOQGNVPSCSYMHGALNNHTQ	963
0Y	618	KSLSLSPG 625	
Db	964	KSLSLSPG 971	

```

1  APPLICANT: Power, Scott D.
2  APPLICANT: Wang, Huaming
3  APPLICANT: Ward, Michael
4  TITLE OF INVENTION: Production of Functional Antibodies in
5  TITLE OF INVENTION: Filamentous Fungi
6  FILE REFERENCE: GC741-2
7  CURRENT APPLICATION NUMBER: US/10/418,836
8  CURRENT FILING DATE: 2003-04-17
9  PRIOR APPLICATION NUMBER: US 60/373,889
10 PRIOR FILING DATE: 2002-04-18
11 PRIOR APPLICATION NUMBER: US 60/411,540
12 PRIOR FILING DATE: 2002-09-18
13 PRIOR APPLICATION NUMBER: US 60/452,134
14 PRIOR FILING DATE: 2003-03-04
15 PRIOR APPLICATION NUMBER: US 60/411,537
16 PRIOR FILING DATE: 2002-09-18
17 NUMBER OF SEQ ID NOS: 40
18 SOFTWARE: FastSeq for Windows Version 4.0
19 SEQ ID NO 39
20 LENGTH: 975
21 TYPE: PRT
22 ORGANISM: Artificial Sequence
23 FEATURE:
24 OTHER INFORMATION: fusion protein
25 US-10-418-836-39

```

Query Match	36.0%	Score 1230.5;	DB 15;	Length 975;
Best Local Similarity	46.7%;	Pred. No. 2.4e-78;		
Matches 284; Conservative	36;	Mismatches 103;	Indels 185;	Gaps 14;

QY	30	LGKGDIVELTCTASQKKSIOF--HMKNQOIKLNGQSFLLKGPBKLANDRDRSRSL-	86
Db	540	LKSESLTSLCTVCSGSLTNYGVHWRQSPGKGLEMIYKWKSGGSTYENAAPISTRLLTIS	599
QY	87	--MOCGFPLIKLKFJESBDTYICEVEDQKEVEQVLVFLGTANSDPHLLQ--GQSLTLT	142
Db	600	KDTSKQNVSLKNSLTAAIDTAIVYC-----ARNDYVANDYMGQGLVLT	642
QY	143	LESPPGSSPSVQCRSPRGKNIQGGKTLVSQLEIODSGTWTCTYVLOKQKVEFKIDIVL	202
Db	643	VSSASTGKPSVFPLAPSPKSTSGG-TAALGCL-----	673
QY	203	AFQCASSIVYKKEEQVEFSPLAFTVTEKLTGSGELMWAERASSKSWITPDLKXKES	262
Db	674	-----VKDYFPEPVYTS-----WNSGALTSG-----	694
QY	263	VKRYTODPKLOMGKKLPLHLTLPOALPOVAGSGLTLALBAKTKGKHQEVNLVYMRATOL	322
Db	695	-----VH-TFPAVL-QSSGLVSSVTVVPSSLSGTQYI-----	727
QY	323	QKMLTCEWGPSTSFKMLMLSLKLEKKAQSKREKPPVAVLNPBAGMQCLLSDSQVLLES	382
Db	728	-----CNV-----NHKFSNTKVDKCV-----	743
QY	383	NIKVLPTWSTPCCPADEPKSCDKTHTC-----PELLGSPSVFLFPPPKPDTLMI	437
Db	744	-----EPKSCDKTHTCCPCPADELLGGSSVFLFPPKPKDTLMI	786
QY	438	VTCAVVVDVSHDEBPVKFNWYVDGVEYHNAKTKPRREQYNSTYRVSVLYVLHODWLNGKE	497
Db	787	VTCAVVVDVSHDEBPVKFNWYVDGVEYHNAKTKPRREQYNSTYRVSVLYVLHODWLNGKE	846
QY	498	YKCAVSNKALPAPLEKTIISAKQGRREPQVYTLPRSPDELTKNOVSLTCLVKGYPSPDIA	557
Db	847	YKCAVSNKALPAPLEKTIISAKQGRREPQVYTLPRSPDELTKNOVSLTCLVKGYPSPDIA	906
QY	558	VEMESNGCPENNYKTPPVLDSDFSFLYSKLTVDKRWMOQGNFSCSVHMEALAHNYTQ	617
Db	907	VEMESNGCPENNYKTPPVLDSDFSFLYSKLTVDKRWMOQGNFSCSVHMEALAHNYTQ	966
QY	618	KSLSLSPG 625	
Db	967	KSLSLSPG 974	


```

QY 441 VVVDVSHEDPEVKFNWYVDCVEYHNAKTKRREOVNSTYRVSVLTVLHODMLNGKEYKC 500
| | | | |
DB 197 VVVDVSHEDPEVKFNWYVDCVEYHNAKTKRREOVNSTYRVSVLTVLHODMLNGKEYKC 256
| | | | |
QY 501 KVSNNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFPYSDIAVEM 560
| | | | |
DB 257 KVSNNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFPYSDIAVEM 316
| | | | |
QY 561 ESNQGEENNYKTPPVLDSDGSFPLYSKLTVDKSRMOQGNVFSQVMEALHNHYTKSL 620
| | | | |
DB 317 ESNQGEENNYKTPPVLDSDGSFPLYSKLTVDKSRMOQGNVFSQVMEALHNHYTKSL 376
| | | | |
QY 621 SLSPG 625
| | | | |
DB 377 SLSPG 381

RESULT 86
US-09-773-877A-18
; Sequence 18, Application US/09773877A
; Publication No. US20030017977A1
; GENERAL INFORMATION:
; APPLICANT: Xia, Yu-Ping et al.
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATORY SKIN DISEASES
; FILE REFERENCE: REG 710b
; CURRENT APPLICATION NUMBER: US/09/773.877A
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Flt1(2-3)-Fc (Muc3)
US-09-773-877A-18

Query Match 36.0%; Score 1229; DB 12; Length 462;
Best Local Similarity 53.9%; Pred. No. 1,1e-78;
Matches 274; Conservative 36; Mismatches 86; Indels 112; Gaps 18;

QY 178 DSGMTCTVLDQ-----NOKKVEFKIDIVLAFQKASIVYKKEGQVEFSPLAF 227
| | | | |
DB 6 DTGVLCAALSLCLLTGSSSGRPVEM-----YSELPELIHMTGR--ELVIPCVR 55
| | | | |
QY 228 TVEKLT-----GSGELMWQAEARASSKSWITFDLKNKEVSXKRVTDPEKIQ 273
| | | | |
DB 56 TSPNITVTLKKFPLDLIPDGKRIW-----DSRKGFIISNATYKEIGL----- 99
| | | | |
QY 274 MGKKLPLHLPLQALPQYASSGNLTALAEKTKL-HQEVNLV-----MRATOLOKN- 325
| | | | |
DB 100 -----LTCEATV-----NGHL-----YKTYNLTRHQNTIIDVQISTPRPVKLLRGH 141
| | | | |
QY 326 ---LTCEVWGPTSPKMLSLKL---ENKEAKVSKR-----EKPVAVLN 362
| | | | |
DB 142 TLVNLCAITATPLNTRVMTSTYDEKAKRASVRRRIQSSSHANIFVSLTIDK---MKN 198
| | | | |
QY 363 PEAGMOCCLSDSGQVLESNIKVLPTWSTPVPAPAPKSCDKTHTC-----PELLGGP 417
| | | | |
DB 199 KDKGLVYCRVR-SGSPFSKSVNTSV-HIYDKAGP---GEPISCDKTHCPCPAPPELLGGP 253
| | | | |
QY 418 SVFLFPKPKDITLMISTPTPEVTCVVVDSHEDPEVKFNWYVDCVEYHNAKTKRREOYNS 477
| | | | |
DB 254 SVFLFPKPKDITLMISTPTPEVTCVVVDSHEDPEVKFNWYVDCVEYHNAKTKRREOYNS 313
| | | | |
QY 478 TYRVSVLTVLHODMLNGKEYKCYSNNKALPAPIEKTISKAKQPREPOVYTLPPSRDEL 537
| | | | |
DB 314 TYRVSVLTVLHODMLNGKEYKCYSNNKALPAPIEKTISKAKQPREPOVYTLPPSRDEL 373
| | | | |
QY 538 TKQOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRMO 597
| | | | |
DB 374 TKQOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRMO 433
| | | | |

```

```

QY 598 OGNVFSQVMEALHNHYTKQSLSLSPG 625
| | | | |
DB 434 OGNVFSQVMEALHNHYTKQSLSLSPG 461
| | | | |

RESULT 87
US-10-433-108-29
; Sequence 29, Application US/10433108
; Publication No. US20040053370A1
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Company
; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
; FILE REFERENCE: X-13991
; CURRENT APPLICATION NUMBER: US/10/433,108
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: US 60/251,954
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-433-108-29

Query Match 36.0%; Score 1228.5; DB 12; Length 272;
Best Local Similarity 87.5%; Pred. No. 5.9e-79;
Matches 231; Conservative 9; Mismatches 13; Indels 11; Gaps 2;

QY 373 SDSGVLESNIKVLPTW-----STPVPCPAPAPKSCDKTHTC-----PELLGGSVFL 421
| | | | |
DB 8 SDSLKQMEEAVALRLFLEWLKNGSPSGAPPSAPKSCDKTHTCPCPAPPELLGGSVFL 67
| | | | |
QY 422 FPKPKDITLMISTPTPEVTCVVVDSHEDPEVKFNWYVDCVEYHNAKTKRREOYNSTYRV 481
| | | | |
DB 68 FPKPKDITLMISTPTPEVTCVVVDSHEDPEVKFNWYVDCVEYHNAKTKRREOYNSTYRV 127
| | | | |
QY 482 VSVLTVLHODMLNGKEYKCYSNNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNQ 541
| | | | |
DB 128 VSVLTVLHODMLNGKEYKCYSNNKALPAPIEKTISKAKQPREPOVYTLPPSRDELTKNQ 187
| | | | |
QY 542 VSVLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRMOQGNV 601
| | | | |
DB 188 VSVLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSFPLYSKLTVDKSRMOQGNV 247
| | | | |
QY 602 FSCSVMEALHNHYTKQSLSLSPG 625
| | | | |
DB 248 FSCSVMEALHNHYTKQSLSLSPG 271
| | | | |

RESULT 88
US-10-404-724-23
; Sequence 23, Application US/10404724
; Publication No. US20030203447A1
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Arnold H.
; TITLE OF INVENTION: Methods and Materials for Increasing Expression of Recombinant
; FILE REFERENCE: 13698U501
; CURRENT APPLICATION NUMBER: US/10/404,724
; PRIOR FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/368,530
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-404-724-23

```

Query Match 36.0%; Score 1228.5; DB 12; Length 465;
 Best Local Similarity 45.8%; Pred. No. 1.2e-78;
 Matches 294; Conservative 35; Mismatches 102; Indels 211; Gaps 17;

```

Qy 11 LVLVQLALPAPATGNNKVLG---KKGDVLTCTASQKKSIOFHMKNNOIKILNOG 66
Db 7 LLLFLMAAQAQAOIQIOLVQSGPEVKKPGESVKISCKAS---GYFTKYGMMVWQAPQGG 63
Qy 67 -----SFLTKGPSKUNDRADSRSLMDQGNF-----LIIKLIKEDSDTYI 108
Db 64 LKMMGWINTYEEPEYGGD-----FKGRFTTLDSTSTAYLLEISLRSDTYTF 114
Qy 109 CEVEDQKEVQLVFLGTLTANSDTHLQGSLLTLESPPGSSPVQCRSPRGNIOGKXT 168
Db 115 C-----ARFGSAVD-----YMGQGLVTVSSASTKGPSVFLPAPSSKSTSGG-T 157
Qy 169 LVSQLELDQSGTWTCTVLQNKVEFKIDIVLAFQKASSIYVKKEGQVEFSPFLAFT 228
Db 158 AALGCL-----WNSGALITSG-----VKDYFPEPYVT 173
Qy 229 VEKLTGSGELMWQABERASSKSWITFDLKNKEVSVKRYTQDPKLGKGLPLHLTLPQAL 288
Db 174 VS-----WNSGALITSG-----VH-TFPAVL 192
Qy 289 PoyAGSNLTLALAKTKGKLHQBENLVVWRATOLQKULTCEVWGPTSPKMLSLKLENKE 348
Db 193 -QSGGLYSLSSVTVPPSSSLGTQTYI-----CNV-----NHK 223
Qy 349 AKVSKREKPVWVLPNPAAGMOCCLSDSGVLLSNIKVLPTWSTPVPCEAPPEKSCDKTH 408
Db 224 PSNTGKVKRV-----EPKSCDKTH 242
Qy 409 TC-----PELLGSPVFLPPPKDQTLMTSRPEVTCVYVDVSHEDPEVKFMYVDGVEV 463
Db 243 TCPCPAPPELLGSPVFLPPPKDQTLMTSRPEVTCVYVDVSHEDPEVKFMYVDGVEV 302
Qy 464 HNAKTKREEOYNSTYRVSVTLVLDHMDLNGKEYKCKVSKALPAPIEKTISKAGOPR 523
Db 303 HNAKTKREEOYNSTYRVSVTLVLDHMDLNGKEYKCKVSKALPAPIEKTISKAGOPR 362
Qy 524 EPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSF 583
Db 363 EPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSF 422
Qy 584 FLVSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
Db 423 FLVSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 464

```

RESULT 89

US-10-656-769-32
 ; Sequence 32, Application US/10656769
 ; Publication No. US2004009712A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Varnum, Brian
 ; APPLICANT: White, Alison
 ; APPLICANT: Vezina, Chris
 ; APPLICANT: Wong, Lu Min
 ; APPLICANT: Qian, Xuejing
 ; TITLE OF INVENTION: Therapeutic Human Anti-IL-1R Monoclonal Antibody
 ; FILE REFERENCE: 01,1554
 ; CURRENT APPLICATION NUMBER: US/10/656,769
 ; PRIORITY FILING DATE: 2003-09-05
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 32
 ; LENGTH: 467
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-656-769-32

Query Match 36.0%; Score 1228; DB 16; Length 467;

Best Local Similarity 45.3%; Pred. No. 1.3e-78;
 Matches 290; Conservative 39; Mismatches 101; Indels 210; Gaps 18;

```

Qy 16 LALLPAAATGQ-----NKVVLGKGDVLTCTASQKKSIOFHM-----KNSN 57
Db 7 LALLLAVLQGVCAEVLQMGAGAVKKKGBESLKISCKGS-GYFSFHMVWWRQMPGKGLE 65
Qy 58 QIKILNGSGFLTKGPS---KLANDRSRSL-----WDQNFPLIINKLIKEDSDTYICE 110
Db 66 WMGIIHFGASDTRSPSPGQVTTISADNSNATYLOV-----SLLKASDTAMFCA 116
Qy 111 VEDQKEVQLVFLGTLTANSDTHLQGSLLTLESPPGSSPVQCRSPRGNIOGKXTLS 170
Db 117 ---RQRELDYFDY-----WQGLTVTVSSASTKGPSVFLPAPSSKSTSGG-TAA 161
Qy 171 VSQLELDQSGTWTCTVLQNKVEFKIDIVLAFQKASSIYVKKEGQVEFSPFLAFTVE 230
Db 162 LGCL-----WNSGALITSG-----VKDYFPEPYVTVS 177
Qy 231 KLTVSGELMWQABERASSKSWITFDLKNKEVSVKRYTQDPKLGKGLPLHLTLPQALPQ 290
Db 178 -----WNSGALITSG-----VH-TFPAVL-Q 195
Qy 291 YAGSNLTLALAKTKGKLHQBENLVVWRATOLQKULTCEVWGPTSPKMLSLKLENKEX 350
Db 196 SSGLYSLSSVTVPPSSSLGTQTYI-----CNV-----NHKPS 227
Qy 351 VSKREKPVWVLPNPAAGMOCCLSDSGVLLSNIKVLPTWSTPVPCEAPPEKSCDKTHTC 410
Db 228 NTGKVKRV-----EPKSCDKTHTC 246
Qy 411 -----PELLGSPVFLPPPKDQTLMTSRPEVTCVYVDVSHEDPEVKFMYVDGVEV 465
Db 247 PRCAPPELLGSPVFLPPPKDQTLMTSRPEVTCVYVDVSHEDPEVKFMYVDGVEV 306
Qy 466 AKTKPREEOYNSTYRVSVTLVLDHMDLNGKEYKCKVSKALPAPIEKTISKAGOPR 525
Db 307 AKTKPREEOYNSTYRVSVTLVLDHMDLNGKEYKCKVSKALPAPIEKTISKAGOPR 366
Qy 526 QVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSF 585
Db 367 QVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEMESNGQPENNYKTPPVLDSDGSF 426
Qy 586 YSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
Db 427 YSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKSLSLSPG 466

```

RESULT 90

US-09-747-669-3
 ; Sequence 3, Application US/09747669
 ; Patent No. US20020122807A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dan, Michael D.
 ; APPLICANT: Saleh, Mansoor
 ; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS, DESIGNATED
 ; TITLE OF INVENTION: 4BS THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES
 ; TITLE OF INVENTION: ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS
 ; FILE REFERENCE: 316082001001
 ; CURRENT APPLICATION NUMBER: US/09/747,669
 ; PRIORITY FILING DATE: 2002-04-08
 ; PRIOR APPLICATION NUMBER: US 09/111,286
 ; PRIORITY FILING DATE: 1998-07-07
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 476
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic construct
 ; US-09-747-669-3

Query Match 36.0%; Score 1228; DB 9; Length 476;
 Best Local Similarity 45.1%; Pred. No. 1,4e-78;
 Matches 289; Conservative 38; Mismatches 110; Indels 204; Gaps 16;

```

16 LALLPAATQGNKV-----LGKKDPEVETCTASQKSIQFH-----52
8 LFLVAATASARQVQLVQSGAEVKKPGASVKVSCKASGVTFTSFDMWRQAPGQGLEMM 67
53 -WNSNQIKILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEY 111
68 GMMNPNPSGK-----TGYAQKFGQRYMTNRTSIRRAY-MELSGLSDEDAVVFCA 117
112 EDQKEEVQL--VFGLTNSDTHLLQGOSLTLTLESPPSSPSVQCRSPRGKNIQSGKTL 169
118 MADNVEMAAIHYHYGND-----VWGQGTIVTSASATKGSVPFLAASSKSTSGG-7A 169
170 SVSOLELQDSGTWCTVLQNKQKVEPKIDIVLAFOKASSIYKKEGEQVEFSFPLAFTV 229
170 ALGCL-----VKGDFPEPPTV 185
230 EKLTSGSELMOAERASSSKSWITPDLKKEVSVKRVTDPKLQMGKULPHLITLPOALP 289
186 S-----WNSGALTSG-----VH-TFPAVL- 203
290 QYAGSGNLTLALEAKTGKLGHEVNLVVMATQLOKVLTCGVWGPTSPKMLSLKLENKA 349
204 QSSGLYSLSVTVTPSSISLGTQTYI-----CNV-----NHKE 235
350 KYSKREKPPVWLNPEAGMOCQLSDSGVLLSNIKVLPTWSTPVCAPAPPEKSCDKHT 409
236 SNTKVDKVV-----EPKSCDKHT 254
410 C-----PELLGSPVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGEVH 464
255 CPKCPAPPELLGSPVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGEVH 314
465 NAKTKREBOYNSTYVSVLTVLHODMLNGEKYKCKVSNKALPAPIEKTISAKQOPRE 524
315 NAKTKREBOYNSTYVSVLTVLHODMLNGEKYKCKVSNKALPAPIEKTISAKQOPRE 374
525 PQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPF 584
375 PQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPF 434
585 LYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
435 LYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 475

```

RESULT 91
 US-10-290-703-3
 ; Sequence 3, Application US/10290703
 ; Publication No. US20030118593A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dan, Michael D.
 ; APPLICANT: Saleh, Mansoor
 ; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS, DESIGNATED
 ; TITLE OF INVENTION: 4B5, THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES
 ; TITLE OF INVENTION: ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS
 ; TITLE OF INVENTION: AND DETECTION OF CANCERS
 ; FILE REFERENCE: 316082001002
 ; CURRENT APPLICATION NUMBER: US/10/290, 703
 ; CURRENT FILING DATE: 2002-11-08
 ; PRIOR APPLICATION NUMBER: US 09/747,669
 ; PRIOR FILING DATE: 2000-12-21
 ; PRIOR APPLICATION NUMBER: US 09/111,286
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: US 60/051,945
 ; PRIOR FILING DATE: 1997-07-08
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3

LENGTH: 476
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic construct
 US-10-290-703-3

Query Match 36.0%; Score 1228; DB 14; Length 476;
 Best Local Similarity 45.1%; Pred. No. 1,4e-78;
 Matches 289; Conservative 38; Mismatches 110; Indels 204; Gaps 16;

```

16 LALLPAATQGNKV-----LGKKDPEVETCTASQKSIQFH-----52
8 LFLVAATASARQVQLVQSGAEVKKPGASVKVSCKASGVTFTSFDMWRQAPGQGLEMM 67
53 -WNSNQIKILGNQGSFLTKGPSKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEY 111
68 GMMNPNPSGK-----TGYAQKFGQRYMTNRTSIRRAY-MELSGLSDEDAVVFCA 117
112 EDQKEEVQL--VFGLTNSDTHLLQGOSLTLTLESPPSSPSVQCRSPRGKNIQSGKTL 169
118 MADNVEMAAIHYHYGND-----VWGQGTIVTSASATKGSVPFLAASSKSTSGG-7A 169
170 SVSOLELQDSGTWCTVLQNKQKVEPKIDIVLAFOKASSIYKKEGEQVEFSFPLAFTV 229
170 ALGCL-----VKGDFPEPPTV 185
230 EKLTSGSELMOAERASSSKSWITPDLKKEVSVKRVTDPKLQMGKULPHLITLPOALP 289
186 S-----WNSGALTSG-----VH-TFPAVL- 203
290 QYAGSGNLTLALEAKTGKLGHEVNLVVMATQLOKVLTCGVWGPTSPKMLSLKLENKA 349
204 QSSGLYSLSVTVTPSSISLGTQTYI-----CNV-----NHKE 235
350 KYSKREKPPVWLNPEAGMOCQLSDSGVLLSNIKVLPTWSTPVCAPAPPEKSCDKHT 409
236 SNTKVDKVV-----EPKSCDKHT 254
410 C-----PELLGSPVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGEVH 464
255 CPKCPAPPELLGSPVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGEVH 314
465 NAKTKREBOYNSTYVSVLTVLHODMLNGEKYKCKVSNKALPAPIEKTISAKQOPRE 524
315 NAKTKREBOYNSTYVSVLTVLHODMLNGEKYKCKVSNKALPAPIEKTISAKQOPRE 374
525 PQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPF 584
375 PQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPVLDSGSPF 434
585 LYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 625
435 LYSKLTVDKSRMOQGNVFCSCVMHEALHNHYTOKSLSLSPG 475

```

RESULT 92
 US-10-433-108-21
 ; Sequence 21, Application US/10433108
 ; Publication No. US20040053370A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eli Lilly and Company
 ; TITLE OF INVENTION: GLP-1 FUSION PROTEINS
 ; FILE REFERENCE: X-13991
 ; CURRENT APPLICATION NUMBER: US/10/433,108
 ; CURRENT FILING DATE: 2003-05-29
 ; PRIOR APPLICATION NUMBER: US 60/251,954
 ; PRIOR FILING DATE: 2000-06-12
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 21
 ; LENGTH: 272
 ; TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
US-10-433-108-21

Query Match 36.0%; Score 1227.5; DB 12; Length 272;
Best Local Similarity 88.3%; Pred. No. 6.9e-79;
Matches 233; Conservative 4; Mismatches 16; Indels 11; Gaps 2;

QY 373 SDSQVLLSNIKVLPTM-----STVPFCAPAPKSCDKTHTC-----PELLGSPSVFL 421
DB 8 SDVSSYLEEQAKAFIALVLKRGSSGAPPPSAEPKSCDKTHTCPPCPAPPELLGSPSVFL 67

QY 422 PPPPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVVHNAKTKPREEOYNSTRV 481
DB 68 PPPPKDTLMISRTPEVTCVVDVSHEDPEVKFMVYDGVVHNAKTKPREEOYNSTRV 127

QY 482 VSVLTVLHQMVLNGEKYCKVSNKALPAPIEKTSKAKGPREPOVYTLPPSRDELTKNQ 541
DB 128 VSVLTVLHQMVLNGEKYCKVSNKALPAPIEKTSKAKGPREPOVYTLPPSRDELTKNQ 187

QY 542 VSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 601
DB 188 VSLTCLVKGFYPSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 247

QY 602 FSCSVMEALHNHYTOKSLSLSPG 625
DB 248 FSCSVMEALHNHYTOKSLSLSPG 271

RESULT 93

US-10-378-567-2
Sequence 2, Application US/10378567
Publication No. US20040006208A1
GENERAL INFORMATION:
APPLICANT: KARPUSAS, MICHAEL
APPLICANT: HSU, YEN-MING
APPLICANT: TAYLOR, FREDERICK R.
APPLICANT: ZHENG, ZHONGLI
TITLE OF INVENTION: CO-CRYSTAL STRUCTURE OF MONOCLONAL ANTIBODY 5C8 AND
TITLE OF INVENTION: CD154, AND USE THEREOF IN DRUG DESIGN
FILE REFERENCE: A096CON1
CURRENT APPLICATION NUMBER: US/10/378.567
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: PCT/US01/27352
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 60/276,452
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/229,933
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 448
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: humanized 5c8 heavy chain amino acid
US-10-378-567-2

Query Match 36.0%; Score 1227.5; DB 15; Length 448;
Best Local Similarity 46.2%; Pred. No. 1.4e-78;
Matches 285; Conservative 42; Mismatches 97; Indels 193; Gaps 17;

QY 25 GNVKLVKGGDTLVLTTASQK--KSIOFHKNISNQLIKLGNQ--SFL-----TKGPSKL 76
DB 8 GAELV--KPGASVLTSCASGIFTSYMYM-----VKQAPGQGLEMIIGELINPENGDTNF 60

QY 77 NDRADSRRLW---DQGNFLLIKNLKIEDSDYIICEVEDQKEEVQLLVFGLTANSPTHL 133
DB 61 NEKFKSKATLVTDKASTAYWELSLRSSEDTAVYVYCTRSRDRNDM----- 106

QY 134 LOGSLTILTESPPGSSPVQCRSPRGNKIQGGKTLVSQLELDOSGTWTCTVLQONQKV 193
DB 107 SMGGTLVTVSSATKPSVFPPLPSSKSTSG--TALGCL----- 146

QY 194 EFKIDIVLAFQKASSIVYKKEGQVEPFPPLAFTVEKLTGSGELMWQABASSKSMIT 253
DB 147 -----VVDYFPEPVTVS-----NNSGALTSG----- 167

QY 254 FDLKNKEVSXRVYQDPKLGKMLPLHLTLPQALPYAGSGNLTALAEATGTLHOENV 313
DB 168 -----VH-TFPAVL--QSSGLYSLSVTVYVSSLSLTQTY 199

QY 314 LVVVRATQIQNLTCFVWGPSTPRLMLSLKENKAVSKREKRVVWVILNPEAGWQCILS 373
DB 200 I-----CNY-----HNKPSNTKVDKV----- 216

QY 374 DSQGVLLSNIKVLPTMSTVPFCAPAPKSCDKTHTC-----PELLGSPSVFLPPPKD 428
DB 217 -----EPKSCDKTHTCPPCPAPPELLGSPSVFLPPPKD 250

QY 429 TLMISRTPEVTCVVDVSHEDPEVKFMVYDGVVHNAKTKPREEOYNSTRVSVTLVL 488
DB 251 TLMISRTPEVTCVVDVSHEDPEVKFMVYDGVVHNAKTKPREEOYNSTRVSVTLVL 310

QY 489 HODVLNGEKYCKVSNKALPAPIEKTSKAKGPREPOVYTLPPSRDELTKNQVSLTCLV 548
DB 311 HODVLNGEKYCKVSNKALPAPIEKTSKAKGPREPOVYTLPPSRDELTKNQVSLTCLV 370

QY 549 KGFPYSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMH 608
DB 371 KGFPYSDIAVWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMH 430

QY 609 EALHNHYTOKSLSLSPG 625
DB 431 EALHNHYTOKSLSLSPG 447

RESULT 94

US-10-108-260A-4285
Sequence 4285, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1 full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108.260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4285
LENGTH: 471
TYPE: PRT
ORGANISM: Homo sapiens
US-10-108-260A-4285

Query Match 35.9%; Score 1226; DB 15; Length 471;
Best Local Similarity 51.2%; Pred. No. 1.8e-78;
Matches 281; Conservative 18; Mismatches 72; Indels 178; Gaps 15;

QY 163 IOGKTLVSQLELDOSGTWTCTVLQONQKVEFKIDIVLAFQKASSIVYKKEGQVEFS 222
DB 14 ISGGQ-----SQVPLVSGT-----EYKKGASVNIIS 40

QY 223 FP--LAFT-----VEKLTGSGELMWQABASSKSMITFDLKNKEVSXRVYQDPKLG 273
DB 41 CKAPGYFTTFYFMHWVWQAPGQGLEMWGIRNPSSGS-----SVSQK----- 82

QY 274 MGKKLPLHLTLPQALPYAGSGNLTALAEATGTLHOENVLVVVRATQIQNLTCF----- 329
DB 83 -----FEGRLTLADISTTANHEL-----RNLTSDDTGV 112

QY 330 -----VMGPSTPKLMLSLKENKAVSKREKRVVWVILNPEA----- 365

Db 113 YCTTRMKVVRGEDNMGQSLVIVSS-----ASTKGSVFLPAPSCKSTSGCT 163
 Qy 366 GMMQCLSD-----SGOVLLSINIKVLP-----TWSTPVPCCAP----- 399
 Db 164 AALGCLVGVYFPEPVTVMWNSG--ALTSGVHTFPAVLQSSGLVLSLSSVTVVPSLSIGTOT 221
 Qy 400 -----EPKSCDKHTTC-----PELLGSPSVFLPPPKDLMISRTPE 436
 Db 222 YICNVNHNKSNVTYVDEKVEPKSCDKHTTCPCAPAPELGGSPSVFLPPPKDLMISRTPE 281
 Qy 437 EYTCVAVDVSHEDPEVKFMVYDGVVHNAKTPREEOYNSTYRVSVLTVLHQMNLNGK 496
 Db 282 EYTCVAVDVSHEDPEVKFMVYDGVVHNAKTPREEOYNSTYRVSVLTVLHQMNLNGK 341
 Qy 497 EYCKKSNKALPAPIEKTISKAGOPREPOVYTLPPSRDELTKNOVSLTCLVGFYPSDI 556
 Db 342 EYCKKSNKALPAPIEKTISKAGOPREPOVYTLPPSRDELTKNOVSLTCLVGFYPSDI 401
 Qy 557 AVEWESNGOPENNYKTTTPVLVDSGSPFLYSKLTVDKSRMQGNVSCSVMEBALNHNHT 616
 Db 402 AVEWESNGOPENNYKTTTPVLVDSGSPFLYSKLTVDKSRMQGNVSCSVMEBALNHNHT 461
 Qy 617 QKSLSLSPG 625
 Db 462 QKSLSLSPG 470

RESULT 95
 US-10-260A-4292
 ; Sequence 4292, Application US/10108260A
 ; Publication No. US20040005560A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA
 ; FILE REFERENCE: H1-A0106
 ; CURRENT APPLICATION NUMBER: US/10/108,260A
 ; CURRENT FILING DATE: 2002-03-27
 ; NUMBER OF SEQ ID NOS: 5458
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 4292
 ; LENGTH: 470
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-108-260A-4292

Query Match 35.9%; Score 1225.5; DB 15; Length 470;
 Best Local Similarity 46.3%; Pred. No. 2e-78;
 Matches 291; Conservative 42; Mismatches 100; Indels 195; Gaps 18;
 Qy 15 QALLLPAAQGNKVVIGKKGDVVELTCTAS--QKKSIOFHMKK-----SNQKIL 62
 Db 20 QVQVQVSGTE---VKKPSSVAVSKASGGSFSTVFTWROAPGEGLEMMGSLTIPIL 74
 Qy 63 GNGGSFLTGPSPKLNDRADSRSLMDQGNFPLIINKLIEDSDTYICEVEDOKEEYQLLV 122
 Db 75 G-RPNVQKQODRTAISADESSSI---VYMDLRLTIEDTAIFYCAI-----LLE 120
 Qy 123 FGLTANSDTLLQGSITLTLESPPGSSPSVQCRSPRGKNIQGGKITLSVQLELDOSGTW 182
 Db 121 HEVVALFD-HMGQGTLTVTSSASTKG--PSVFPLAPSKSTSGG-TAALGCL----- 168
 Qy 183 TCTVTLQNGKVEFRIDIVLAFQKASIVYKKEGEQVEFSPLAFTEYKLTGSGELMWQA 242
 Db 169 -----VKDYFPEPVTVS-----WNS 183
 Qy 243 ERASSSKSWITFDLKNKEVSKVKTODPKLQMGKPLHLTLPOALPOVAGSGLTVALE 302
 Db 184 GALTSG-----VH-TFPAVL-QSSGIVSLSSVVT 210
 Qy 303 AKTGKLGHOENVLVVMRATOLQKNLTCEVWGPSTPKMLSLKENKAIVSKREKPVWVLN 362
 Db 211 VPSSISLOTQTYI-----CNV-----NHNKPSNTKVDKTV----- 238

Qy 363 PEAGMMQCLSDSGOVLLSINIKVLPWTSTPVPCCAPPEPKSCDKHTTC-----PELLGSP 417
 Db 239 -----EPKSCDKHTTCPCAPPELLGSP 261
 Qy 418 SVFLPPPKDLMISRTPEVTCVAVDVSHEDPEVKFMVYDGVVHNAKTPREEOYN 477
 Db 262 SVFLPPPKDLMISRTPEVTCVAVDVSHEDPEVKFMVYDGVVHNAKTPREEOYN 321
 Qy 478 TYRVSVLTVLHQMNLNGKYEYCKKSNKALPAPIEKTISKAGOPREPOVYTLPPSRDEL 537
 Db 322 TYRVSVLTVLHQMNLNGKYEYCKKSNKALPAPIEKTISKAGOPREPOVYTLPPSRDEL 381
 Qy 538 TKNOVSLTCLVGFYPSDI AVEWESNGOPENNYKTTTPVLVDSGSPFLYSKLTVDKSRMQ 597
 Db 382 TKNOVSLTCLVGFYPSDI AVEWESNGOPENNYKTTTPVLVDSGSPFLYSKLTVDKSRMQ 441
 Qy 598 QGNVSCSVMEBALNHNHTQKSLSLSPG 625
 Db 442 QGNVSCSVMEBALNHNHTQKSLSLSPG 469

RESULT 96
 US-10-282-162-34
 ; Sequence 34, Application US/10282162
 ; Publication No. US20030143697A1
 ; GENERAL INFORMATION:
 ; APPLICANT: REGENERON PHARMACEUTICALS, INC.
 ; TITLE OF INVENTION: RECEPTOR BASED ANTAGONISTS, AND METHODS OF MAKING
 ; FILE REFERENCE: REG 203-B-US
 ; CURRENT APPLICATION NUMBER: US/10/282,162
 ; CURRENT FILING DATE: 2002-10-28
 ; PRIOR APPLICATION NUMBER: 09/787,835
 ; PRIOR FILING DATE: 1999-09-22
 ; PRIOR APPLICATION NUMBER: PCT/US99/22045
 ; PRIOR FILING DATE: 1999-09-22
 ; NUMBER OF SEQ ID NOS: 56
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 34
 ; LENGTH: 900
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-282-162-34

Query Match 35.9%; Score 1225.5; DB 14; Length 900;
 Best Local Similarity 48.3%; Pred. No. 4.8e-78;
 Matches 284; Conservative 45; Mismatches 112; Indels 147; Gaps 19;
 Qy 86 LW-----DOGNFPLIINKLIEDSDTYICEVEDOKEEYQLLV-FVGLTANSDTLLH--Q 135
 Db 411 LWFRTLLNDTGNITQMLN-----TYCSKVAFLLEVQKDSCTNSPMKLPVHKLTYIE 464
 Qy 136 GOSLTLTLESPPGSSPS-----VQCRSPRGKN-IQGGKITLSVQLELDOSGTWC 184
 Db 465 YGIORTCPNVGDYFPSSVAKPTITWYMGCYKIQNFNNVPIEGNLSFLALISNNGYTC 524
 Qy 185 TVLQNGKVEF---KIDIVLAFQKAS-----SIYVKE-GEQ-----VEFSF 223
 Db 525 VVTPYENGRTFHLTRLTLTVKVGSPKNAVAPVHISPNHVHVEKEGEELLIPCTYFESF 584
 Qy 224 PLAFTEKLTGSGELMWQERASSSKSWITFDLK-NKEVSKVKTODPKLQMGKPLHL 282
 Db 585 -----LWDSRHEVMTID--GKKDDITIDVTINESISHKTBETRQI----- 627
 Qy 283 TLPOALPOVAGSGLTVALEAKTKLHOENVLVVMRATOLQKNLTCEVWGPSTPKMLSL 342
 Db 628 -----LSIKVTSB-----DLKRSYVCHA----- 646
 Qy 343 KLENKAIVSKREKPVWVLNPEAGMMQCLSDSGOVLLSINIKVLPWTSTPVPCCAPPEPK 402
 Db 647 ---RSK-----GEVAKAKVK-----QKVAPRYTVE 671
 Qy 403 SCDKHTTC-----PELLGSPSVFLPPPKDLMISRTPEVTCVAVDVSHEDPEVKFMV 457

Db 672 SGDTHHTPCPCAPABELLGGPSVFLPPPKDITLMSRPEVTVVVDVSHDEPEKFWY 731
Qy 458 VDGVEVHNAKTRPEEQVNSTYRVSVLTVHOMLNKEKCKRSNALPAPIEKTISK 517
Db 732 VDGVEVHNAKTRPEEQVNSTYRVSVLTVHOMLNKEKCKRSNALPAPIEKTISK 791
Qy 518 AKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVWESNGOPENNYKTPPV 577
Db 792 AKGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVWESNGOPENNYKTPPV 851
Qy 578 DSDGSFFLYSKLTVDKSRMOQGNVSCSVMEHALNHYTOKSLSPG 625
Db 852 DSDGSFFLYSKLTVDKSRMOQGNVSCSVMEHALNHYTOKSLSPG 899

RESULT 97
US-10-683-255-12
; Sequence 12, Application US/10683255
; Publication No. US20040063910A1
; GENERAL INFORMATION:
; APPLICANT: Kavanaugh, William M.
; APPLICANT: Ballinger, Marcus
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR
; TITLE OF INVENTION: RECEPTOR-IMMUNOGLOBULIN FUSION
; FILE REFERENCE: PP01474.101
; CURRENT APPLICATION NUMBER: US/10/683,255
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 09/499,846
; PRIOR FILING DATE: 2000-02-07
; PRIOR APPLICATION NUMBER: 60/119,002
; PRIOR FILING DATE: 1999-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-683-255-12

Query Match 35.9%; Score 1225; DB 12; Length 488;
Best Local Similarity 67.3%; Pred. No. 2.3e-78;
Matches 253; Conservative 18; Mismatches 56; Indels 49; Gaps 8;
Qy 293 GSGULTALEAKTGKTHQ---EVNLVYMRATQLOKTL--TCEWGPSPKMLSLKE-- 345
Db 118 GSINHTYQLDVERSPHRPILOAGLPANKTVALGSNVEFMCKVYSDPQPHIQWLKHIEVN 177
Qy 346 -----NKEAKVSKREKPVWVNL-----PEAGMOCILSDSGQVLLSENIK 385
Db 178 GSKTGPDLRYVQILKTAGVTTDKEMEVLHLRNVSPEDAGEYTCLAGNSIGLSHNS--- 234
Qy 386 VLPWTST-----PVPCPAP---EPKSCDKTHTC-----PELLGGPSVFLPPPKDXT 429
Db 235 ---AWLTVALERPAVMTSPVLEPSCDKTHTCPCPAPBELGGSVLFPKPKDXT 291
Qy 430 LMIKRTPEVTVVVDVSHDEPEVKFNMYVDGEVHNAKTRPEEQVNSTYRVSVLTVH 489
Db 292 LMIKRTPEVTVVVDVSHDEPEVKFNMYVDGEVHNAKTRPEEQVNSTYRVSVLTVH 351
Qy 490 ODMNGKREYKCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLTCLV 549
Db 352 ODMNGKREYKCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLTCLV 411
Qy 550 GFYPSDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVSCSVME 609
Db 412 GFYPSDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVSCSVME 471
Qy 610 ALNHNHYTOKSLSPG 625
Db 472 ALNHNHYTOKSLSPG 487

RESULT 98
US-10-193-616-14
; Sequence 14, Application US/10193616
; Publication No. US20030096355A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Ke
; TITLE OF INVENTION: Isolation, Identification, and Characterization of
; TITLE OF INVENTION: Ymk5, a novel
; TITLE OF INVENTION: member of the TNF-Receptor Supergene Family
; FILE REFERENCE: 01017/33551A
; CURRENT APPLICATION NUMBER: US/10/193,616
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/611,989
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/143,137
; PRIOR FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: ymk5-FC fusion protein
US-10-193-616-14

Query Match 35.9%; Score 1224.5; DB 14; Length 396;
Best Local Similarity 92.1%; Pred. No. 1.9e-78;
Matches 233; Conservative 1; Mismatches 14; Indels 5; Gaps 1;
Qy 378 VLESNIKVLPTWSTPVPCPAPPEKSCDKTHTC-----PELLGGPSVFLPPPKDITLMI 432
Db 143 VLQECNSTANTYVCSSSNMAAPEKSCDKTHTCPCPAPBELGGPSVFLPPPKDITLMI 202
Qy 433 SRTEPEVTVVVDVSHDEPEVKFNMYVDGEVHNAKTRPEEQVNSTYRVSVLTVHQM 492
Db 203 SRTEPEVTVVVDVSHDEPEVKFNMYVDGEVHNAKTRPEEQVNSTYRVSVLTVHQM 262
Qy 493 LNKREYKCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVGFY 552
Db 263 LNKREYKCKVSNKALPAPIEKTISKAKGPREPOVYTLPPSRDELTKNOVSLTCLVGFY 322
Qy 553 PSDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVSCSVMEHALH 612
Db 323 PSDIAVWESNGOPENNYKTPPVLDSDGSFFLYSKLTVDKSRMOQGNVSCSVMEHALH 382
Qy 613 NHYTOKSLSPG 625
Db 383 NHYTOKSLSPG 395

RESULT 99
US-10-471-151-32
; Sequence 32, Application US/10471151
; Publication No. US20040086908A1
; GENERAL INFORMATION:
; APPLICANT: Chandrasekhar, Yasmin A.
; APPLICANT: Novak, Julia E.
; APPLICANT: Foster, Donald C.
; APPLICANT: Wenfeng, Xu
; APPLICANT: Jaspers, Stephen R.
; TITLE OF INVENTION: Soluble Heterodimeric Cytokine Receptor
; FILE REFERENCE: 01-10PC
; CURRENT APPLICATION NUMBER: US/10/471,151
; PRIOR FILING DATE: 2003-09-08
; PRIOR APPLICATION NUMBER: 60/274,560
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/299,865
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 541

TYPE: PRT
ORGANISM: Homo sapiens
US-10-471-151-32

Query Match 35.9%; Score 1224.5; DB 16; Length 541;
Best Local Similarity 53.8%; Pred. No. 2.8e-78;
Matches 278; Conservative 26; Mismatches 100; Indels 113; Gaps 13;

```

QY 209 SIYKKEGEQ-----VEFSPLAFTYBKLTGSGELMMQARASSSKSWITFDLK 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 37 SIEKTYGERDVAKKGCCORITRKSCNLTETGNLT---ELYARVTAASAGRSATKMT 93
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 258 NKEVSVKRYTQDP-----KLQMGKULPLHLTPQALPOYAGSGNLTALAEKTK 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 94 DRPSSLGHTTLKPPDYTCISKVRSIQM-----IVHPTPTIRAGDG-RLTLEIDIFHD 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 308 LHOENVLVVVRATOLQK-----LTCEVWGPTSPKMLSLK-L 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 146 LFYHLEQVNRRTYQMHGKQREYFEGLTPTDEFLGTIMICVPTAKSAPYMCVKTL 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 345 ENKEAKVSKREKPVWVNLPEA-----GMQCLLSD-----SGVLLBSNINIKY 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 206 PDRTWTASTKGPSVFPLAPSSKSTSGGTALGCLVXDYFPEPVTVSMNSG--ALTSGVHT 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 387 LP-----TWSTPVPCEAP-----EPKSCDKTHTC--- 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 264 FPAVLQSSGLYSLSSVTVVPSSSLGTQTYICNVNHRKSTKVDKKEPSPSCDKTHTCPPC 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 --PELLGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKT 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 324 PAPELLGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKT 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 469 KPREQYNSTYRVSVLTTLHODWLNKGEYKCKVSNKALPAPIEKTIISAKQGPREFQYV 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 384 KPREQYNSTYRVSVLTTLHODWLNKGEYKCKVSNKALPAPIEKTIISAKQGPREFQYV 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 529 TLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNQPENNYKTTTPVLDSDGSFELYSK 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 444 TLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNQPENNYKTTTPVLDSDGSFELYSK 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 589 LTVDKSRWQGNVFCSCVMEALHNHYTQKSLSLSPG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 504 LTVDKSRWQGNVFCSCVMEALHNHYTQKSLSLSPG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 100
US-10-471-151-31
Sequence 31, Application US/10471151
Publication No. US20040086908A1
GENERAL INFORMATION:
APPLICANT: Chandrasekhar, Yashmin A.
APPLICANT: Novak, Julia E.
APPLICANT: Foster, Donald C.
APPLICANT: Wenfeng, Xu
APPLICANT: Jaspers, Stephen R.
TITLE OF INVENTION: Soluble Heterodimeric Cytokine Receptor
FILE REFERENCE: 01-10PC
CURRENT APPLICATION NUMBER: US/10/471,151
PRIOR FILING DATE: 2003-09-08
PRIOR APPLICATION NUMBER: 60/274,560
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/299,865
PRIOR FILING DATE: 2001-06-21
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 31
LENGTH: 558
TYPE: PRT
ORGANISM: Homo sapiens
US-10-471-151-31

Query Match 35.9%; Score 1224.5; DB 16; Length 558;
Best Local Similarity 53.8%; Pred. No. 3e-78;

```

Matches 278; Conservative 26; Mismatches 100; Indels 113; Gaps 13;
QY 209 SIYKKEGEQ-----VEFSPLAFTYBKLTGSGELMMQARASSSKSWITFDLK 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 54 SIEKTYGERDVAKKGCCORITRKSCNLTETGNLT---ELYARVTAASAGRSATKMT 110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 258 NKEVSVKRYTQDP-----KLQMGKULPLHLTPQALPOYAGSGNLTALAEKTK 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 111 DRPSSLGHTTLKPPDYTCISKVRSIQM-----IVHPTPTIRAGDG-RLTLEIDIFHD 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 308 LHOENVLVVVRATOLQK-----LTCEVWGPTSPKMLSLK-L 344
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 163 LFYHLEQVNRRTYQMHGKQREYFEGLTPTDEFLGTIMICVPTAKSAPYMCVKTL 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 345 ENKEAKVSKREKPVWVNLPEA-----GMQCLLSD-----SGVLLBSNINIKY 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 223 PDRTWTASTKGPSVFPLAPSSKSTSGGTALGCLVXDYFPEPVTVSMNSG--ALTSGVHT 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 387 LP-----TWSTPVPCEAP-----EPKSCDKTHTC--- 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 281 FPAVLQSSGLYSLSSVTVVPSSSLGTQTYICNVNHRKSTKVDKKEPSPSCDKTHTCPPC 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 --PELLGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKT 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 341 PAPELLGSPVFLPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVNAKT 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 469 KPREQYNSTYRVSVLTTLHODWLNKGEYKCKVSNKALPAPIEKTIISAKQGPREFQYV 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 401 KPREQYNSTYRVSVLTTLHODWLNKGEYKCKVSNKALPAPIEKTIISAKQGPREFQYV 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 529 TLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNQPENNYKTTTPVLDSDGSFELYSK 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 TLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNQPENNYKTTTPVLDSDGSFELYSK 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 589 LTVDKSRWQGNVFCSCVMEALHNHYTQKSLSLSPG 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 LTVDKSRWQGNVFCSCVMEALHNHYTQKSLSLSPG 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: August 3, 2004, 13:48:02
Job time : 67.8475 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 13:00:24 ; Search time 5.92937 Seconds
(without alignments)
4185.504 Million cell updates/sec

Title: SEQ6
Perfect score: 1317
Sequence: 1 MNREVPRHLLVQLALP.....VISFLGIGVACVLRTR 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1023	77.7	458	1 RWHUT4	T-cell surface gly
2	875	66.4	432	1 RWCZT4	T-cell surface gly
3	790	60.0	432	1 RWCQT4	T-cell surface gly
4	596.5	45.3	459	2 A46254	CD4 precursor - ra
5	562	42.7	432	2 S30193	T-cell surface gly
6	480.5	37.2	457	2 A27449	T-cell surface gly
7	475	36.1	457	1 RWMST4	T-cell surface gly
8	357	27.1	71	1 I60082	CD4 receptor - hum
9	332	25.2	240	2 A39016	T-cell surface gly
10	305.5	23.2	99	2 S21461	T-cell surface gly
11	280.5	21.3	99	2 S21462	T-cell surface gly
12	156.5	11.9	538	2 JC2457	vascular cell adhe
13	133	10.1	1259	2 S36126	neural cell adhesi
14	131	9.9	1260	1 S05479	neural cell adhesi
15	129	9.8	338	2 JC4776	limbic-system-asso
16	124.5	9.5	584	2 I50419	8-glycerin precurs
17	124	9.4	739	2 J08675	vascular cell adhe
18	123	9.3	338	2 J05519	50K glycoprotein p
19	120	9.1	1091	2 A58532	glial cell membran
20	120	9.1	1277	2 T30532	neural cell adhesi
21	120	9.1	2222	2 T13924	sdh protein - frui
22	118.5	9.0	584	2 T08678	hypothetical prote
23	118	9.0	1091	2 S01958	contractin precurs
24	118	9.0	1197	2 T30581	neural cell adhesi
25	118	9.0	6831	2 A88852	protein unc-22 (im
26	118	9.0	6839	2 S57242	twitichin [smilari
27	118	9.0	727935	2 T27935	hypothetical prote
28	117.5	8.9	345	2 S03199	opioid-binding pro
29	117	8.9	739	2 A41288	vascular cell adhe

30	117	8.9	1011	2 T13669	neuromusculin - fr
31	116.5	8.8	345	2 JC4025	opioid-binding cel
32	115	8.7	647	2 B41288	vascular cell adhe
33	113	8.6	1018	2 JC4211	neural adhesion pr
34	112.5	8.5	773	1 QRRBG	secretory componen
35	111.5	8.5	120	2 S46374	ig kappa chain V-J
36	111	8.4	702	2 A36319	carcinoembryonic a
37	111	8.4	1091	2 S33850	fibronectin-binding
38	110.5	8.4	398	2 I49443	gene 2B4 protein -
39	110.5	8.4	2029	1 TDFLK	protein-tyrosine-p
40	110	8.4	521	2 S34338	biliary glycoprote
41	110	8.4	739	2 JN0581	vascular cell adhe
42	109.5	8.3	279	2 S04693	T-cell receptor de
43	109.5	8.3	304	2 S04653	T-cell receptor ga
44	109	8.3	1257	1 A41060	neural cell adhesi
45	109	8.3	1367	2 A41228	protein-tyrosine k
46	108.5	8.2	333	2 A31923	amalgam protein pr
47	108	8.2	257	2 S00682	IGF Fc receptor al
48	107.5	8.2	122	2 S40370	ig kappa chain - h
49	107	8.1	761	1 TJHNG	neural cell adhesi
50	107	8.1	3707	2 S18252	heparan sulfate pr
51	107	8.1	4391	2 A38096	perlecan precursor
52	106.5	8.1	111	2 B37266	ig kappa chain V r
53	106.5	8.1	111	2 I38740	ig kappa chain V r
54	106.5	8.1	345	2 JC1239	opioid-binding pro
55	106.5	8.1	2783	2 T34416	hypothetical prote
56	106	8.0	1232	2 T43027	neural cell adhesi
57	106	8.0	1896	2 T08851	Down syndrome cell
58	105.5	8.0	210	2 I49294	CD7 antigen - mous
59	105.5	8.0	1323	2 P00568	connectin 3B - chi
60	105	8.0	2629	2 T32735	telomerase-associat
61	104.5	7.9	1239	1 A32579	neuroglian - fruit
62	104	7.9	338	2 JC1238	opioid-binding pro
63	104	7.9	344	2 I56551	neurotrophin - rat
64	104	7.9	458	2 JC1509	biliary glycoprote
65	104	7.9	1018	2 A54744	contractin 1 precu
66	104	7.9	1020	2 S05944	neural cell surf
67	104	7.9	1021	2 A57112	contractin precurs
68	103.5	7.9	103	2 S18731	ig kappa chain V-J
69	103.5	7.9	108	1 RVM506	ig kappa chain V r
70	103.5	7.9	117	2 S21668	ig kappa chain V r
71	103.5	7.9	129	1 K1HWK	ig kappa chain pre
72	103.5	7.9	1028	2 A53449	plasmacytoma-asso
73	103.5	7.9	5175	2 T20992	hypothetical prote
74	103.5	7.9	5198	2 T43280	neutricin precurs
75	103	7.8	519	2 A44783	ecto-ATPase precu
76	102.5	7.8	862	2 I49583	differentiation an
77	102.5	7.8	1348	2 S51656	vascular endotheli
78	102	7.7	458	1 WMM5R1	biliary glycoprote
79	102	7.7	458	2 S23969	cell-adhesion mole
80	102	7.7	458	2 S68177	C-CAM2a protein is
81	102	7.7	521	2 JC1508	biliary glycoprote
82	102	7.7	7962	2 I38346	elastic titin - hu
83	101.5	7.7	98	2 PH1066	ig light chain V r
84	101.5	7.7	725	1 TJMSNG	neural cell adhesi
85	101.5	7.7	858	1 IURMNC	neural cell adhesi
86	101.5	7.7	1115	1 TJMSNL	neural cell adhesi
87	101	7.7	135	2 S22383	hypothetical prote
88	101	7.7	1036	2 S22383	axonin 1 precursor
89	101	7.7	3375	2 T19821	hypothetical prote
90	100.5	7.6	108	1 RVM561	ig kappa chain V r
91	100.5	7.6	765	2 C42632	cell adhesion mole
92	100.5	7.6	806	2 A35963	protein-tyrosine k
93	100.5	7.6	812	2 B42632	cell adhesion mole
94	100.5	7.6	932	2 A42632	cell adhesion mole
95	100.5	7.6	26926	1 I38344	titin, cardiac mus
96	100	7.6	464	2 C30127	transmembrane carc
97	100	7.6	526	1 A32164	biliary glycoprote
98	99	7.5	230	2 S33161	ig kappa chain - s
99	98.5	7.5	129	2 S40317	ig kappa chain - h
100	98.5	7.5	129	2 A34815	carcinoembryonic a
101	98.5	7.5	1028	2 I58164	BIG-1 protein - ra
102	98	7.4	483	2 T17346	hypothetical prote

```

103 98 7.4 764 1 ORHUGS secretory component
104 98 7.4 880 2 B53743 protein-tyrosine k
105 97.5 7.4 108 1 KWS582 Ig kappa chain V r
106 97.5 7.4 117 2 S11700 Ig kappa chain pre
107 97.5 7.4 122 2 S40314 Ig kappa chain - h
108 97.5 7.4 620 2 JH0593 Schwann cell myel
109 97.5 7.4 1333 2 T78875 receptor tyrosine
110 97.5 7.4 6642 2 T29757 protein UNC-89 - C
111 97 7.4 1394 2 T29925 hypothetical prote
112 97 7.4 470 2 S52080 Ig heavy chain pre
113 97 7.4 729 2 A56795 fibroblast growth
114 97 7.4 733 2 I49293 fibroblast growth
115 97 7.4 832 2 I49289 fibroblast growth
116 97 7.4 1091 1 IJCHNL neural cell adhesi
117 97 7.4 1328 2 T23007 hypothetical prote
118 97 7.4 1330 2 S49010 embryonic receptor
119 97 7.4 1447 2 A54100 tumor suppressor p
120 96.5 7.3 117 1 K1H012 Ig kappa chain pre
121 96.5 7.3 122 2 S40351 Ig kappa chain V-J
122 96.5 7.3 125 2 S40349 Ig kappa chain V-J
123 96.5 7.3 130 2 PL0113 Ig kappa chain pre
124 96.5 7.3 323 2 S01895 T-cell receptor ga
125 96.5 7.3 365 2 JC7780 coxsackie- and ade

```

ALIGNMENTS

```

RESULT 1
RMHUT4
T-cell surface glycoprotein CD4 precursor [validated] - human
N:Alternate names: T-cell surface antigen T4/Len 3
C:Species: Homo sapiens (man)
C:Date: 20-May-1986 #sequence, revision 31-Dec-1988 #text, change 20-Apr-2001
C:Accession: A90872; A32722; A34194; A53287; I54176; I54297; A02109; A30039
R:Maddon, P.J.; Littman, D.R.; Godfrey, M.; Maddon, D.E.; Chess, L.; Axel, R.
Cell 42, 93-104, 1985
A:title: The isolation and nucleotide sequence of a cDNA encoding the T cell surface pr
A:Reference number: A90872; MUID:85254946; PMID:2990730
A:Accession: A90872
A:Molecule type: mRNA
A:Residues: 1-25, 'N', 27-458 <MAD>
A:Experimental source: clone p748
R:Littman, D.R.; Maddon, P.J.; Axel, R.
Cell 55, 541, 1988
A:title: Corrected CD4 sequence.
A:Reference number: A90907; MUID:89028665; PMID:3263213
A:Contents: annotation; revision to residue 26
R:Camerini, D.; Seed, B.
Cell 60, 747-754, 1990
A:title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi
A:Reference number: A32722; MUID:90182664; PMID:2107024
A:Accession: A32722
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 26-426, 428-458 <CAM>
R:Carri, S.A.; Hemling, M.E.; Folema-Wasserman, G.; Sweet, R.W.; Annunzio, K.; Barr, J.R.;
J. Biol. Chem. 264, 21286-21295, 1989
A:title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 recep
A:Reference number: A34194; MUID:90078232; PMID:2592374
A:Contents: disulfide bonds; carbohydrate-binding sites
A:Accession: A34194
A:Molecule type: protein
A:Residues: 26-394 <CAR>
R:Lederman, S.; Demartino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, A
Mol. Immunol. 28, 1171-1181, 1991
A:title: A single amino acid substitution in a common African allele of the CD4 molecule
A:Reference number: A53287; MUID:92072595; PMID:1961196
A:Accession: A53287
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 250-264, 'W', 266-280 <LED>
A>Note: sequence extracted from NCBI backbone (NCBIF:68249)

```

```

R:Edwards, M.C.; Gibbs, R.A.
Genomics 14, 590-597, 1992
A:title: A human dimerism resulting from loss of an Alu.
A:Reference number: I54176; MUID:93052387; PMID:1330888
A:Accession: I54176
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-72 <RES>
A:Cross-references: GB:U47924; GB:M6525; GB:U72506; NID:g1633547; PIDN:AA51309.1; PID
Hum. Immunol. 30, 99-104, 1991
A:title: Humans with OKT4-epitope deficiency have a single nucleotide base change in th
A:Reference number: I54297; MUID:91216786; PMID:1708753
A:Accession: I54297
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-264, 'W', 266-458 <RE2>
A:Cross-references: GB:M35160; NID:g179143; PIDN:AA16069.1; PID:g179144
F:421-458/Domain: Intracellular #status predicted <INT>
F:421-109, 155-184, 328-370/Disulfide bonds: #status experimental
F:596, 325/Binding site: carbohydrate (Aen) (covalent) #status experimental
Query Match 77.7%; Score 1023; DB 1; Length 458;
Best Local Similarity 99.5%; Pred. No. 2, 6e-67;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MNRGVPRHLLVLTQLALPAATQGNKVVYGGKGDVVELTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVLTQLALPAATQGNKVVYGGKGDVVELTCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNFPLIKNLKIESDPTTYICEVEQKEVOL 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNFPLIKNLKIESDPTTYICEVEQKEVOL 120
QY 121 LVFGLTRANSPTHLQGOQLTLTLESPGSSPVQCRSPKXNQQGKTLVSQLELDPSG 180
DB 121 LVFGLTRANSPTHLQGOQLTLTLESPGSSPVQCRSPKXNQQGKTLVSQLELDPSG 180
QY 181 TWICTVLQNOQKVEFKIDIV 200
DB 181 TWICTVLQNOQKVEFKIDIV 200
QY 181 TWICTVLQNOQKVEFKIDIV 200
DB 181 TWICTVLQNOQKVEFKIDIV 200
RESULT 2
RMCTZ4
T-cell surface glycoprotein CD4 - chimpanzee
N:Alternate names: T-cell surface antigen T4/Len 3
C:Species: Pan troglodytes (chimpanzee)
C:Date: 30-Sep-1993 #sequence, revision 30-Sep-1993 #text, change 16-Jul-1999
R:Camerini, D.; Seed, B.
Cell 60, 747-754, 1990
A:title: A CD4 domain important for HIV-mediated syncytium formation lies outside the v
A:Reference number: A32722; MUID:90182664; PMID:2107024
A:Accession: B32722
A:Molecule type: mRNA
A:Residues: 1-432 <CAM>
A:Cross-references: GB:M31135

```

R:/Pomgaard, A.; Hirsch, V.M.; Johnson, P.R.

Eur. J. Immunol. 22, 2973-2981, 1992

A/Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor

A/Reference number: A46534; MUID:93049640; PMID:1442921

A/Accession: A46534

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 3-399 <FOM>

A/Note: sequence extracted from NCBI backbone (NCBIF:118332)

C/Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells

C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

C/Keywords: duplication; glycoprotein; T-cell; transmembrane protein

F/1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>

F/1-371/Domains: extracellular #status predicted <EXT>

F/9-86/Domains: immunoglobulin homology <IM1>

F/111-161/Domains: immunoglobulin homology #status atypical <IM2>

F/191-274/Domains: immunoglobulin homology <IM4>

F/296-347/Domains: immunoglobulin homology <IM3>

F/372-395/Domains: transmembrane #status predicted <TM>

F/396-432/Domains: intracellular #status predicted <INT>

F/16-84,130-159,303-345/Disulfide bonds: #status predicted

F/271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 66.4%; Score 875; DB 1; Length 432;

Best Local Similarity 97.7%; Pred. No. 1.5e-56;

Matches 170; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 27 KVLVGGKDDVTELTCTASOKKSIQFHMKNNOIKILNQGSFLTKGPKLNDRAISRSL 86

DB 2 KVLVGGKDDVTELTCTASOKKSIQFHMKNNOIKILNQGSFLTKGPKLNDRAISRSL 61

QY 87 WDQGNFPLIINKLIEDSDTYICEVDQKEEVQLVFGLTANSPTHLLOQSILTLTLESP 146

DB 62 WDQGNFPLIINKLIEDSDTYICEVDQKEEVQLVFGLTANSPTHLLOQSILTLTLESP 121

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

QY 27 KVLVGGKDDVTELTCTASOKKSIQFHMKNNOIKILNQGSFLTKGPKLNDRAISRSL 86

DB 2 KVLVGGKDDVTELTCTASOKKSIQFHMKNNOIKILNQGSFLTKGPKLNDRAISRSL 61

QY 87 WDQGNFPLIINKLIEDSDTYICEVDQKEEVQLVFGLTANSPTHLLOQSILTLTLESP 146

DB 62 WDQGNFPLIINKLIEDSDTYICEVDQKEEVQLVFGLTANSPTHLLOQSILTLTLESP 121

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175

QY 147 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 200

DB 122 PGSSPSVQCRSPRKNIQGGKTLVSQLELDQSGTWCTVLOQNKVPEFKIDIV 175


```

QY      1  MNRGVPFRH-LLVLVQLAL-PPATQGNKVLGKKGDFVELTCTASQOKKSIOFHKNKSNQI 59
      1  MCRALISRLILLLLLOSLAVTQGTVLVKGESSEALFPCSSQKKITVFTKFSQDR 60
      60  KILGNQGSFPLTKG--PSKLNDRADSRSLMDQGNFPLIINKLTIKEDSDTYICEVEDQKE 116
      61  KILGQHGKGVLRIGGSPSQF--DRFDSKKGAWEKGSFPLIINKLAMEDSGQTYICELENRKE 119
QY      117  EVOLLVGLTRANSTHLLQGSLLTLTLFS--PRGSSPSVQCRSPRGKXIOGKTLISVQGLE 175
      120  EVELMVFVKVTFSPGTSLLQGSLLTLTLDSNKSXSNPLTECKGKXGVSGSKVLSMSNLR 179
QY      176  LODSGTWTCTVLONQKK 192
      180  VQSDPFNNCTVTLDOCK 196
Db

RESULT 8
160082
CD4 receptor - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 23-Jul-1999
C/Accession: 160082
R/Zverev, V.V.; Sidorov, A.V.; Nedospasov, S.A.; Malinuhova, V.V.; Udaloova, I.A.; Andzha
Vopr. Virusol. 40, 100-102, 1995
A/Title: [Nucleotide sequence of two exons of the human T-lymphocyte CD4 receptor gene].
A/Reference number: 160082; MUID:95407135; PMID:7676667
A/Accession: 160082
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-71 <RES>
A/Cross-references: GB:579267; NID:g1086922; PIDN:AAB35273.1; PID:g1086923
C/Genetics:
A/Introns: 17/1
C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology

Query Match      27.1%; Score 357; DB 2; Length 71;
Best Local Similarity 96.6%; Pred. No. 1.2e-19;
Matches 70; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  MNRGVPFRHLLVLVQLAL-PPATQGNKVLGKKGDFVELTCTASQOKKSIOFHKNKSNQI 60
      1  MNRGVPFRHLLVLVQLAL-PPATQGNKVLGKKGDFVELTCTASQOKKSIOFHKNKSNQI 60
Db
QY      61  ILGNQGSFLTK 71
      61  ILGNQGSFLTK 71
Db

RESULT 9
A39016
T-cell surface glycoprotein CD7 precursor - human
N/Alternate names: T-cell leukemia antigen
C/Species: Homo sapiens (man)
C/Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 05-Nov-1999
C/Accession: A39016; S03520
R/Schambery, L.E.; Fleenor, D.E.; Kurtzberg, J.; Haynes, B.F.; Kaufman, R.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 603-607, 1991
A/Title: Isolation and characterization of the genomic human CD7 gene: structural simila
A/Reference number: A39016; MUID:91110576; PMID:1703303
A/Accession: A39016
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-240 <SCH>
A/Cross-references: GB:M37271; NID:g180163; PIDN:AAA51953.1; PID:g180164
R/Aruffo, A.; Seed, B.
EMBO J. 6, 3313-3316, 1987
A/Title: Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a COS cell expr
A/Reference number: S03520; MUID:88111517; PMID:3501369
A/Molecule type: mRNA
A/Residues: 1-240 <ARU>

```

```

A/Cross-references: EMBL:X06180; NID:g29819; PIDN:CAA29546.1; PID:g732757
C/Genetics:
A/Gene: GDB:CD7
A/Cross-references: GDB:119770; OMIM:186820
A/Map position: 17q25.2-17q25.3
A/Intons: 28/1
C/Keywords: glycoprotein; membrane protein; surface antigen; tandem repeat
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-240/Product: T-cell surface glycoprotein CD7 #status predicted <MNT>
F:145-180/Region: 9-residue repeats (X-P-P-X-A-S-A-L-P)

Query Match      25.2%; Score 332; DB 2; Length 240;
Best Local Similarity 36.4%; Pred. No. 3.5e-17;
Matches 102; Conservative 18; Mismatches 54; Indels 106; Gaps 10;

QY      6  PFRHLLVLQAL--LPATQGNKVLGKK-----GDVELTCTASQOKKSIOFHKNKSN 56
      4  PFRHLLVLQALARGALAAQVQGSFHCCTVPVGAASVNTICSTSG----- 52
Db
QY      57  NQIKILGNQGSFPLTK-GPS-----KLNDRADSRSLMDQGNFPLIINK 98
      53  -----GKGIYLRQLQEPQODIITYEDGVVPTTDRFRGRIDPSGS---QDNLITIMHR 103
QY      99  LKIEDSDTYICEVEDQKEVQLVFGILTANSTHLLQGSLLTLTLSPGSSPSVQCRSP 158
      104  LQSDTGTYNC-----QATTEV----- 120
Db
QY      159  RGNKIOGKTLVSQLELQDSGTWTCTVLONQKKEFKDIVPAPALPAPPTGSALPDP 218
      121  ---NVYSGSLT-VLVTEEQSGQWHRCS-----DAPPASALPAPPTGSALPDP 164
      219  QPASALPDPAPASALPAPALVIFSLGLGVACVLAART 258
      165  QPASALPDPAPASALPAPALVIFSLGLGVACVLAARTQ 204
Db

RESULT 10
S21461
T-cell surface glycoprotein CD4 (allele 1) - pig (fragment)
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 21-Jan-2000
C/Accession: 147131; S21461
R/Gustafsson, K.; Germana, S.; Sundt, T.M.
J. Immunol. 151, 1365-1370, 1993
A/Title: Extensive allelic polymorphism in the CDR2-like region of the miniature swine
A/Reference number: 147131; MUID:93329116; PMID:8335933
A/Accession: 147131
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-99 <GU2>
A/Cross-references: EMBL:X65629; NID:g1928; PIDN:CAA46583.1; PID:g388232
C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C/Keywords: glycoprotein; T-cell
F:3-81/Domain: immunoglobulin homology <IMM>

Query Match      23.2%; Score 305.5; DB 2; Length 99;
Best Local Similarity 60.2%; Pred. No. 1.1e-15;
Matches 59; Conservative 16; Mismatches 22; Indels 1; Gaps 1;

QY      32  KKGDTVELTCTASQOKKSIOFHKNKSNQIKILGNQGSFL--TKGPKLNDRADSRSLMDQ 90
      1  KAGDLAEVLPCHSSQOKKLPFNKNSNOTKILGHSFMTASTVELTSLRDSKKNMWDHG 60
Db
QY      91  NPLIINKLTIKEDSDTYICEVEDQKEVQLVFGILTAN 128
      61  SFLIINKLTVTDSGYICEVEDKRIEVLQVLRPLTAS 98
Db

RESULT 11
S21462
T-cell surface glycoprotein CD4 (allele 2) - pig (fragment)
C/Species: Sus scrofa domestica (domestic pig)
C/Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 21-Jan-2000

```

C:\Accession: 147132; S21462
R:\Guertafson, K.; Germana, S.; Sundt, T.M.
J. Immunol. 151, 1365-1370, 1993
A>Title: Extensive allelic polymorphism in the CDR2-like region of the miniature swine C
A:Reference number: 147131; MUID:9332916; PMID:8335933
A:Accession: 147132
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-99 <GU>
A:Cross-references: EMBL:X65630; NID:g1929; PIDN:CAA46584.1; PID:g388233
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology
C:Keywords: glycoprotein; T-cell
F:3-81/Domain: immunoglobulin homology <IM>

Query Match 21.3%; Score 280.5; DB 2; Length 99;
Best Local Similarity 56.1%; Pred. No. 7.le-14;
Matches 55; Conservative 19; Mismatches 23; Indels 1; Gaps 1;

Db

32 KKGDVETLTCAASQKSIQFHWNKNNOIKILNGSGFLTKGP-SKLNDRADSRRLMDQG 90
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
1 KAGLAELPCHSSSQKNDLPFSWKNSDDKILRSRRNLHMKAVTESSLSDSKKNWDHG 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |

Qy 91 NFPLIINKLKIEDSDTYICEVEDQKEEYQLTFGLTAN 128
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
61 SFPLIINKLEVTDSGIYICEVEDKRIVQLVFPLTMS 98
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 12

JC2457
vascular cell adhesion protein - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 15-Feb-1995 #sequence_revision 05-Apr-1995 #text_change 08-Oct-1999
C:Accession: JC2457
R:\Tang, Y.T.M.; Haskard, D.O.; Robinson, M.K.
Biochem. Biophys. Res. Commun. 201, 805-812, 1994
A>Title: Cloning and expression kinetics of porcine vascular cell adhesion molecule-
A:Reference number: JC2457; MUID:94271236; PMID:7516159
A:Accession: JC2457
A:Molecule type: mRNA
A:Residues: 1-538 <TSA>
A:Cross-references: EMBL:U06351; NID:G474382; PIDN:AA21542.1; PID:G474383
C:Keywords: glycoprotein; transmembrane protein
F:497-517/Domain: transmembrane #status predicted <TM>
F:75,157,271,330,360/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.9%; Score 156.5; DB 2; Length 538;
Best Local Similarity 24.5%; Pred. No. 0.0006;
Matches 49; Conservative 39; Mismatches 71; Indels 41; Gaps 6;

Qy 32 KKGDVETLTCAASQKSIQFHWNKNNOIKILNGSGFLTKGPSKLNDRADSRRLMDQN 91
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 235 QEGDSMMWTCSTSEGLPAPOISM-----SKLDNGDQQLT---SGN 271
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Qy 92 FPLIINKLKIEDSDTYICE----VEDQKEEYQLLV-----FGLTANSDFHLGGOSLT 141
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 272 ATLTIARMEDSGIYVEGVNPVTNNKEVELTVQAAPDTTISVPSSTLREGSSVM 331
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Qy 142 TLSPSPGSSPV---QCSPRGKNIQGGKTLVSQLODSGTWTVLQ---NQKV 193
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 332 TCSSDGFPAPRIKMSKLRDGNLEPLSENTLTILTSTMEDSGIYVEGINQAINKEY 391
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Qy 194 EFKIDIVPASALPAPTGS 213
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 392 ELIIQAPKDQLQITAFPSSES 411
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 13

S36126
neural cell adhesion molecule L1 - rat
N:\Alternate names: nerve growth factor-inducible large external glycoprotein; NILE glycc
C:Species: Rattus norvegicus (Norway rat)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
C:Accession: S36126; S17655; A60917; A30326

R:Miura, M.; Kobayashi, M.; Asou, H.; Uyemura, K.
FEBS Lett. 289, 91-95, 1991
A:Title: Molecular cloning of cDNA encoding the rat neural cell adhesion molecule L1. T
A:Reference number: S17655; MUID:91372414; PMID:1894011
A:Accession: S36126
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1259 <MIU>
A:Cross-references: EMBL:X59149
A:Accession: S17655
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178, 1183-1259 <MI2>
R:Prince, J.T.; Milona, N.; Stallcup, W.B.
J. Neurosci. 9, 1825-1834, 1989
A:Title: Characterization of a partial cDNA clone for the N1E glycoprotein and identifi
A:Reference number: A60917; MUID:89257627; PMID:2733751
A:Accession: A60917
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1159-1199, 'G', 1201-1235, 'K', 1237 <PRI>
A:Note: this paper appeared earlier, with printing errors, as reference A30326
R:Prince, J.T.; Milona, N.; Stallcup, W.B.
J. Neurosci. 9, 876-883, 1989
A:Title: Characterization of a partial cDNA clone for the N1E glycoprotein and identifi
A:Reference number: A30326; MUID:89177485; PMID:2466966
A:Contents: annotation
A:Note: this paper was reprinted as reference A60917 to correct the omission of several
C:Comment: This sequence of this surface-accessible glycoprotein differs at only two po
accessible only after treatment of cells with detergent and is assumed to be cytoplasm
C:Superfamily: neural cell adhesion molecule L1, fibronectin type III repeat homology;
C:Keywords: cell adhesion; duplication; glycoprotein; membrane protein
F:J531-592/Domains: immunoglobulin homology <IMW>

Query Match 10.1%; Score 133; DB 2; Length 1259;
Best Local Similarity 24.4%; Pred. No. 0.086;
Matches 57; Conservative 31; Mismatches 96; Indels 50; Gaps 9;

QY 10 LLLVLIQDALLPAAATQGNKVVLGKKGDVVELTCTASQKKSID-FHKNNSNQIKILGNQGS 67
DB 507 ILAVIQVKAATQIQGPRSTIEKKGARVTFQGSFSPBSLQASITWRGDR----- 557
QY 68 FLTGPSPKLRADRSKSLMDQGNFPLIKLKIKEDSDTYIC---EVEDQKEVQLLVF 123
DB 558 -----DLQERGSGDKVFIEDGQ-LVTKSLDSDQGVCAVSTELDEVESRAQLLV 608
QY 124 GLTAN-----SDPHLLQGGSLTLTLESPPSSSPSVQCRRSP-----RGKNIQGGKTL 169
DB 609 GSPGPVPHLEISDHLKQSQVHLN-----SFAEDHNSIEKTDIEFEDKNAPEKMF 662
QY 170 SVSOLLEDSGTWTCTVQLQNOQKVE--FKDIVRASALPAPPTGSALPBDQA 221
DB 663 SLGKV---PENGQSTTLKLSPVVHYHFRVAIKKYGGBSPVSETVVTVEAA 712

RESULT 14
S05479
neural cell adhesion molecule L1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S05479; B60850; S22167
R:Moos, M.; Tacke, R.; Scherter, H.; Teplow, D.; Frueh, K.; Schachner, M.
Nature 334, 701-703, 1988
A:Title: Neural adhesion molecule L1 as a member of the immunoglobulin superfamily with
A:Reference number: S05479; MUID:88318924; PMID:3412448
A:Accession: S05479
A:Molecule type: mRNA
A:Residues: 1-1260 <MOO>
A:Cross-references: EMBL:X12875; NID:953336; PIDN:CAA31368.1; PID:953337
A:Note: the authors translated the codon CCG for residue 166 as Leu, ACT for residue 39
R:Rathjen, F.G.; Wolff, J.M.; Frank, R.; Bomhoeffer, F.; Rutishauser, U.

J. Cell Biol. 104, 343-353, 1987

A:Title: Membrane glycoproteins involved in neurite fasciculation.

A:Reference number: A60850; MUID:87109457; PMID:3805123

A:Accession: B60850

A:Molecule type: protein

A:Residues: 20-28, 'XX', 31-36 <RAT>

R:Kohl, A.; Gliese, K.P.; Mohajeri, M.H.; Montag, D.; Moos, M.; Schachner, M.

submitted to the EMBL Data Library, December 1991

A:Description: Analysis of promoter activity and 5' genomic structure of the neural cellin

A:Reference number: S22167

A:Accession: S22167

A:Molecule type: DNA

A:Residues: 1-165, 'L', 167-169, 'E', 191-281, 'S', 283-395, 'S', 397-514, 'APEKNPDV', 524, 'GGNNE

A:Cross-references: EMBL:X63511

C:Genetics:

A:introns: 26/1; 31/1; 66/2; 133/1; 174/1; 231/1; 268/2; 330/1; 374/1; 422/1; 459/2

A:Note: The list of introns may be incomplete

C:Keywords: alternative splicing; cell adhesion; cell adhesion; duplication; glycoprotein; transmembran

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1260/Product: neural cell adhesion molecule #status experimental <MAT>

F:256-313/Domain: immunoglobulin homology <IMM1>

F:440-498/Domain: immunoglobulin homology <IMM2>

F:531-592/Domain: immunoglobulin homology <IMM3>

Query Match 9.8%; Score 131; DB 1; Length 1260;

Best Local Similarity 23.9%; Pred. No. 0.12;

Matches 56; Conservative 33; Mismatches 95; Indels 50; Gaps 9;

Qy 10 LLLVQLALLPRAQGNKVLGKKGDVLELTCTASQKSID-PRHKNNSNIIKLNGQS 67

Db 507 LLNLQVEATQIQPRSALEKGAARVTFCTQASFPDSLOASTWGDGR----- 557

Qy 68 FLTKGPSKTLNDRASRRSLMDQGNFPLIKMLKIEDSPTYIC---EVEDKEEVLVLF 123

Db 558 -----DLGRGSDSKFIEDGK--LVIGSLDSDGNGYSCVASTELDEVESAQLVLF 608

Qy 124 GLTAN-----SDTHLLQGQSLTLTLESPPSSPSVOCRSP-----RCKNIQGKTL 169

Db 609 GSPGPVPHLESLDRHLTKQSQVHLMSW-----SPAEDHNSPIEKYDIEFEDKEMAPKWF 662

Qy 170 SVSQLELDSDGTWCTVYLGNQKYE--FKDIVPRASLPRPGSALPDPQTA 221

Db 663 SLGRV---PCNGQSTTLTKLSPYVHYTFRTAINKYGGEPSPYSESVTPEEA 712

RESULT 15

JC4776

1 limbic-system-associated membrane protein precursor - human

C:Species: Homo sapiens (man)

C:Date: 10-May-1996 #sequence_revision 16-Aug-1996 #text_change 19-May-2000

C:Accession: JC4776

R:Pimenta, A.F.; Fischer, I.; Levitt, P.

Gene 170, 189-195, 1996

A:Title: cDNA cloning and structural analysis of the human limbic-system-associated mem

A:Reference number: JC4776; MUID:96235133; PMID:8666243

A:Accession: JC4776

A:Molecule type: mRNA

A:Residues: 1-338 <PIM>

A:Cross-references: GB:A19191; NID:g12766898; PIDN:AAC50569.1; PID:g1276899

A:Experimental source: brain

C:Comment: This is a neuronal surface glycoprotein distributed in cortical and subcortical

C:Genetics:

A:Gene: lamp

C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi

C:Keywords: brain; glycoprotein; membrane protein; phosphoprotein

F:1-7/Domain: signal sequence #status predicted <SIG>

F:333-338/Region: hydrophobic

F:40,66,136,148,279,287,300,315/Binding site: carboxydrate (Asn) (covalent) #status pred

F:42,115,142,164,171,220,231/Binding site: phosphate (Thr) (covalent) #status predicted

F:95,197,204,226,310/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 9.8%; Score 129; DB 2; Length 338;

```

Best Local Similarity 23.2%; Pred No. 0.035;
Matches 57; Conservative 40; Mismatches 91; Indels 56; Gaps 11;

Qy      10 LILVQLALLPPA-----TQGNKRVLGKKGDTVELTCTASOKKSIOFHKNNSNQIKI 61
       |||:|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      14 LVLLALLCLLPGLFVRSVDVFNRGTDNITVRGGDTALNLCVLEDKNS-KVAMLNLSGILF 72

Qy      62 LGNGGSFLTKGPSKLNDRA--DSRRSLMDQGNFPPLIIKNLKIEDSDTYICEVDQKE--- 116
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      73 AGHD-----KMSLDPRVELEKRHSL----EYSLRIQKVDYDESGSYCYOTQHPEPKT 121

Qy      117 -EVOLLVG---LRANSTPHLLQGOSLTLTLESPPGSSPSVOCR--SPRKNIOGGKT- 168
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      122 SGVVLIIVQVPKRNISISSDVTVNEGSNVTLVCMANGREPPIVTWRHLPTGRFEGBEEY 181

Qy      169 LSVSOLELODSSGTWTCVLQ-----NOKKVEFKIDIVP-----RA 203
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      162 LEILDITHEOGSKRECKAKANEVSADVKQKTVAVYPPIITSEKSNENATTTGQASLKCA 241

Qy      204 SALPAP 209
       |||::|||
Db      242 SAVPAP 247

RESULT 16
150419
s-glycerin precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: IS0419
R:Taira, E.; Takana, N.; Taniura, H.; Kim, C.H.; Mihi, N.
Neuron 12, 861-872, 1994
A>Title: Molecular cloning and functional expression of glycerln, a novel cell adhesion
A:Reference number: I50419; MUID:94213753; PMID:8161457
A:Accession: IS0419
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-584 <TAI>
A:Cross-references: GB:DJ8559; NID:g1009246; PIDN:BA07563.1; PID:g559701

Query Match          9.5%; Score 124.5; DB 2; Length 584;
Beet Local Similarity 23.7%; Pred.No.0.14; 88; Indels 43; Gaps 6;
Matches 50; Conservative 30; Mismatches

Qy      19 LPAAIQGNKVVLGKKGDVTELCTASOKKSIOFHKNNSNQIKILNGGGSFLTKGPSKLND 78
       |||:|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      268 LRVAINAGIV---KSGDVKLVCAD-----GNPAVFSPFRRELQD 306

Qy      79 RADSRSLMDQGNFPPLIIKNLKIEDSDTYICEVD-----OKEEVOLIV-----FEL 125
       |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      307 SWODWTSLADTNVDGLMHNWSKSSSGLYRCQTLLDMTGHEGVLEVNVYIEGVQVMK 366

Qy      126 TRANSTHLLQGOSLTLTLESPPGSSPSVOGRSPRKNIQGGKTLTSVSQLELODSSGWTC 185
       ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      367 EPSSPLH--EGSVRLSCTASHPVKLDYQWRARARKVAEGNQLLTNLTFETSSNFSCR 424

Qy      186 V-----LONOKKVEFKIDIVRASALPAP 209
       |||::|||::|||::|||::|||::|||::|||
Db      425 VKARSVPGLEBSKOVAVAVKGKPRIVAISAP 455

RESULT 17
JS0675
vascular cell adhesion molecule-1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 05-Nov-1999
C:Accession: JS0675; S19872; S23136
R:Hesslein, C.; Moy, P.; Tizard, R.; Chisholm, P.; Williams, C.; Wyck, M.; Burkly, L.; M
Biochem. Biophys. Res. Commun. 183, 163-169, 1992
A>Title: Cloning of murine and rat vascular cell adhesion molecule-1.
A:Reference number: JS0674; MUID:92181437; PMID:1371918
A:Accession: JS0675
A>Status: nucleic acid sequence not shown
```


A: Molecule type: mRNA
A: Residues: 1-739 <HES>
A: Cross-references: GB:M84488, NID:g207642, PIDN:AAA42332.1, PID:g207643
R:Williams, A.; Atkins, R.; Fries, J.; Gimbrone, M.A.; Cybulsky, M.I.; Collins, T.
Submitted to the EMBL Data Library, February 1992
A: Description: Nucleotide sequence of rat vascular cell adhesion molecule-1.
A: Reference number: S19872
A: Accession: S19872
A: Molecule type: mRNA
A: Residues: 1-2, 'G', 4-121, 'HL', 124-165, 'N', 167-738, 'G', <WIL>
A: Cross-references: EMBL:X63722, NID:957471, PIDN:CAA45254.1, PID:957472
R:Williams, A.J.; Atkins, R.C.; Fries, J.W.U.; Gimbrone Jr., M.A.; Cybulsky, M.I.; Collins, T.
Biochim. Biophys. Acta 1131, 214-216, 1992
A: Title: Nucleotide sequence of rat vascular cell adhesion molecule-1 cDNA.
A: Reference number: S23136, MUID:92305064, PMID:1377031
A: Accession: S23136
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-2, 'G', 4-165, 'N', 167-738, 'G', <WIL>
C: Comment: This protein interacts with the beta-1 integrin very late antigen 4 on leukocytes.
A: Gene: VCAM-1
C: Superfamily: immunoglobulin homology
C: Keywords: cell adhesion; transmembrane protein
F: 1-24/Domain: signal sequence #status predicted <SIG>
F: 25-739/Product: vascular cell adhesion molecule 1 #status predicted <VAS>
F: 259-293/Domain: immunoglobulin homology <IMM1>
F: 328-385/Domain: immunoglobulin homology <IMM2>
F: 526-581/Domain: immunoglobulin homology <IMM3>
F: 676-696/Domain: transmembrane #status predicted <TRA>
F: 697-715/Domain: intracellular #status predicted <INT>

Query Match 9.4%; Score 124, DB 2, Length 739;
Best Local Similarity 23.4%; Pred. No. 0.21;
Matches 43; Conservative 30; Mismatches 71; Indels 40; Gaps 5;

33 KGDVLTCTAQQKSIQPHMKNQKILGNQSFITKPSKLNDRADRSIMDQNF 92
Db 238 EGAAVNTTCASGEGPAPEITWSSK-----LDNGVLQL-----SGNA 274
Qy 93 PLIIKNIKIEDSDTYICE-----VEDQKEVQLV-----FGLTANSDTHLLOGQSILTT 142
Db 275 TLTLIARMEDSGIYVEGVNLVGRDKTEVELIYQKRFPTVDISPGQVAAGVDSVLTT 334
Qy 143 LESPSSSPVQVCSPPGKNIQ-----GKITSVQLEIHDGSGTCTVTLONQKVER 195
Db 335 CAANGCSPFSWRTQTDSPLNGEVRDEGATSTLTLSPVGEDEHSYLCVTCCORRLEK 394
Qy 196 KIRI 199
Db 395 TIQV 398
RESULT 18
JCS519
50K glycoprotein precursor - chicken
C: Species: Gallus gallus (chicken)
C: Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 17-Nov-2000
C: Accession: JCS519
R: Hancock, K.A.; Gooley, A.A.; Jeffrey, P.L.
Mol. Brain Res. 44, 273-285, 1997
A: Title: AVGP50, a predominantly axonally expressed glycoprotein, is a member of the Igl
A: Reference number: JCS519, MUID:97225899, PMID:9073169
A: Accession: JCS519
A: Molecule type: mRNA
A: Residues: 1-338 <HAN>
A: Experimental source: brain
C: Comment: This protein belongs to the IGLON's subfamily of cell adhesion molecules.
C: Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
C: Keyword: glycoprotein
F: 1-31/Domain: signal sequence #status predicted <SIG>
F: 32-338/Product: 50K glycoprotein #status predicted <MAT>
F: 46-113/Domain: immunoglobulin homology <IMM1>

F: 146-199/Domain: immunoglobulin homology <IMM2>
F: 232-292/Domain: immunoglobulin homology <IMM3>
F: 40, 136, 148, 279, 300, 315/Binding site: carbohydrate (Asn) (covalent) #status predic
Query Match 9.3%; Score 123, DB 2, Length 338;
Best Local Similarity 24.1%; Pred. No. 0.097;
Matches 59; Conservative 41; Mismatches 107; Indels 38; Gaps 11;

Qy 10 LLIIVLQALLPAA-----TQGNKVLGKKGDVLTCTASQKSIQPHMKNQIKI 61
Db 14 LVTLRLICLLPGLPYRVDFTFGDINIVRQGDITLILCFVDRSS-KVAMNRSGLIF 72
Qy 62 LGNGSFILTKGPSKLNDRAD-SRRSLMDQNFILIKNIKIEDSDTYICEVEDQ-----KE 116
Db 73 AG-----EDKSLDPKRVLEKRSPL--YSLRIQKVDVDEGSYTCGVQTHHPKTS 122
Qy 117 EVQLVFG-----LTANSDTHLLOGQSILTTLESPSSSPVQCR--SPRKNIQGKT-L 169
Db 123 QVYLIVQVPKISNISDITVNGSVTLVCMANGRPPEVITRHLTLPGKEFGESEVYL 182
Qy 170 SVSQLELDGSGTCTVQLQ----NOKVPEKIDIV--RASALPAPPTGSALPQQTAS 222
Db 183 ELIGITREOSGKYECKRANVASADVQVRYVTVNYPPTTESKSNPAATGQALLPCEAS 242
Qy 223 ALPDP 227
Db 243 AVFTP 247

RESULT 19
A58532
glial cell membrane glycoprotein LIG-1 precursor - mouse

C: Species: Mus musculus (house mouse)
C: Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999
C: Accession: A58532
R: Suzuki, Y.; Sato, N.; Tohyama, M.; Manaka, A.; Takagi, T.
J. Biol. Chem. 271, 22522-22527, 1996
A: Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in
A: Reference number: A58532, MUID:96394313, PMID:8798419
A: Accession: A58532
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-1091 <SUZ>
A: Cross-references: GB:D19872, NID:g1545806, PIDN:BA111416.1, PID:g1545807
C: Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-te
F: 36-61/Domain: proteoglycan amino-terminal homology <PAR>
F: 71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F: 95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F: 142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F: 166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F: 191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F: 214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F: 238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F: 262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F: 286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F: 310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F: 334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
F: 358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
F: 385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
F: 409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
F: 440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 9.1%; Score 120, DB 2, Length 1091;
Best Local Similarity 22.3%; Pred. No. 0.64;
Matches 60; Conservative 40; Mismatches 103; Indels 66; Gaps 10;

Qy 14 LQIALPANTQGNKVVLGKKGDVLTCTASQKSIQPHMKNQKILGNQSFILTKGP 73
Db 595 LTVNVLDSFTKLTGHDIAIRGTTARLECAATGPNPQIAKQKGG-----TDFF 643
Qy 74 SKLNDRADRSIMDQNFPLIIKNIKIEDSDTYICEVEDQKEVQLVGLTANSDTHL 133

Db 644 A-----ARERRHVMPPDDVFETIDVKIDMGVSVCTAONSAGSV-----SANATLTV 691
Oy 134 LOGGSLTLTLSS-----PPGSSPVOCR-----SPRGKIOGGKTLSS----- 170
Db 692 LETSLSAVPLEDRVYVTEVAFAQKATGSPPTRTWKGGPPLSLTERHHFTPGNOQLV 751
Oy 171 VSOLELDSGTWTCTVLONOKKVEFKIDIVPRA--SALPAPPGSALPPDPOTASALPDPPA 229
Db 752 VQNMWDAGRYTCE--MSNPGLTE-----RAHSQSLILPTPGCKKGTIVGIF----- 798
Oy 230 ASALPALAVTSFLGLGLGAVACVILATRR 258
Db 799 -----TIAVVCISIVLTSLVWVCIIYQTR 821

RESULT 20
T30532
neural cell adhesion molecule L1 homolog - Fugu rubripes
C/Species: Fugu rubripes
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Mar-2002
C/Accession: T30532
R:Riboldi Tunicliffe, G.R.; Platzter, M.; Nakamura, G.; Elgar, G.S.; Brenner, S.; Rosen
submitted to the EMBL Data Library, September 1997
A/Description: Analysis of the genomic loci of Fugu rubripes homologs of the human disease
A/Reference number: 220848
A/Accession: T30532
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1277 <RIB>
A/Cross-references: EMBL:AF026198; NID:g3098263; PID:g3098264; PIDN:AAC15860.1
C/Genetics:
A/Intons: 42/1; 47/1; 81/2; 149/1; 190/1; 247/1; 285/2; 347/1; 391/1; 440/1; 477/2; 531/2
A/Note: L1-CAM
C/Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; 1

Query Match 9.1%; Score 120; DB 2; Length 1277;
Best Local Similarity 24.7%; Pred. No. 0.78;
Matches 55; Conservative 34; Mismatches 92; Indels 42; Gaps 9;

Oy 34 GDTVELTCTASQKSIQFHWKNSNOIKILNQGSFLTGPRLKNDRADSRSLMDQGNFP 93
Db 363 GEYVRLDQADGIPSPRTW-----TVNGVPLSATSLSPRSLTRESGS-- 405
Oy 94 LIINKLIKEDSDTYICEVEDQKE-----VQLVFGITANSDFH--LLOGSLTL-- 141
Db 406 LILMDVIFGDPAIVQCAQNSNGTILANTNVYIELPQILTEGNTYTFVEGGQKALLEC 465
Oy 142 -TLESP-----PGSSPVOCRSPRGKIOGGKTLSSVSOLELDSGTWTCTVLONOKKVE 194
Db 466 ETFSPPKRYTWESSSISLLADPRVNLITNG--GLEIANVSHDEGIIYTCLVQGSNINSVN 524
Oy 195 FKIDIVPRASALPAPPGSALPDPPA-----SALPDPPASAL 233
Db 525 AEVEVLNRTVTL--SPQALRLQPKKTAIFCLVYTDPLSSPL 566

RESULT 21
T13924
sdc protease - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C/Accession: T13924
R:Nguyen, D.N.; Liu, Y.; Litcky, M.L.; Reinke, R.
submitted to the EMBL Data Library, February 1997
A/Description: Silexick, a member of the immunoglobulin superfamily, is required for pat
A/Reference number: 217809
A/Accession: T13924
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-2222 <NGU>
A/Cross-references: EMBL:U80578; NID:g4099554; PID:g4099555; PIDN:AAD09632.1
C/Genetics:

A/Gene: sdc
A/Cross-references: FlyBase:FBgn0021764

Query Match 9.1%; Score 120; DB 2; Length 2222;
Best Local Similarity 24.0%; Pred. No. 1.5;
Matches 54; Conservative 30; Mismatches 85; Indels 56; Gaps 11;

Oy 26 NKVYLGGKGDVELTCTASQKSIQFHWKNSNOIKILNQGSFLTGPRLKNDRADSRSL 85
Db 463 NVTALDGDADTISCRANGSPNPNTWYNETQLVDI-----SSIVQ 503
Oy 86 LMDQGNPPLIINKLIKEDSDTYIC-----EVEDQKEVQLVFGTL--ANSDFHLOGQ 137
Db 504 ILSEGD--LILSNIRSVADAPLYICVRANVAGSVAAEAVLSVYRTQIIQPPVDTVLIG- 560
Oy 138 SLTLTLSSPPGSSSV-----QCRSPRGK-----IQGKTLSSVSOLELDSGTWT 183
Db 561 -LTATTLQCKVSSDSPYVNIIDWREGQSTPISNSORIGVQADQLEIQAVRASDVGSYA 619
Oy 184 CTVAQ-----NOKKVEFKIDIVPRASALPAPPG--SALPDPPAS 222
Db 620 CVVTSPPGNETRAA--RLSVI-----ELPPPSNVKVERLPEQQA 659

RESULT 22
T08678
hypothetical protein DKFZp56411922.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
C/Accession: T08678
R:Mambur, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A/Reference number: Z16469
A/Accession: T08678
A/Molecule type: mRNA
A/Residues: 1-584 <MAN>
A/Cross-references: EMBL:AL049946
A/Experimental source: fetal brain; clone DKFZp56411922
C/Genetics:
A/Note: DKFZp56411922.1

Query Match 9.0%; Score 118.5; DB 2; Length 584;
Best Local Similarity 21.2%; Pred. No. 0.4;
Matches 59; Conservative 38; Mismatches 90; Indels 91; Gaps 12;

Oy 32 KKGDTVELTC-----TASQKSIQFHWKNSNOIKILNQGSFLT-----KGPRLKND-R 79
Db 165 QRSDSGNVYTCVLRNSAGEDRKTVI--VNVQPPKINGNPMTTVREIAGGSRKLIDCK 223
Oy 80 AD---SRRSLMD-----QGNPPLIINKLIKEDSDTYICEVEDQKEEV 118
Db 224 AEGIPTRVLMAPPEGVLPAPYVGNRIYVNGSGSLDIRLSKSDSVQLCMANEGGEA 283
Oy 119 QLT-----VFGITANSDFHLOGQSLTLTLESPGSSPVOCRSPRGKIOGG 166
Db 284 RLIIQLTVLEPMKPRTHDISEKITMAAGHTISLNSAAGTTPPSLVWVLPKNTDIQSG 343
Oy 167 KTLSS-----VSOLELDSGTWTCTVLON-----QKVEFKIDIVPRAS----- 204
Db 344 QQLQRFYHKADGMLHISGLSSVDAGAVRC--VARNAAGHTRLVSLKVLKPEASKQYHNL 402
Oy 205 -----ALPAPPGSA-----LPDPPT 220
Db 403 VSIINGETLKLCTPPGAGQGRFSWTLPLNGMHLEGPOT 440

RESULT 23
S01998
contactin precursor - chicken
N/Alternate names: 130K glycoprotein
C/Species: Gallus gallus (chicken)
C/Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 17-Nov-2000
C/Accession: S01998; J00094

N:Contains: protein kinase (EC 2.7.1.-)
 C:Species: Caenorhabditis elegans
 C>Date: 28-Oct-1995 #sequence_revision 24-Oct-1997 #text_change 20-Jun-2000
 C:Accession: S57242; S07571; S06797; S57218; T27934; T28030
 R:Benian, G.M.; L'Hernault, S.W.; Morris, M.E.
 Submitted to the EMBL Data Library, February 1993
 A:Description: Additional sequence complexity within twitching of Caenorhabditis elegans
 A:Reference number: S57242
 A:Accession: S57242
 A:Molecule type: DNA
 A:Residues: 1-6839 <BEN1>
 A:Cross-references: EMBL:L10351
 A:Experimental source: var. Bristol
 R:Benian, G.
 Submitted to the EMBL Data Library, November 1989
 A:Reference number: S07571
 A:Accession: S07571
 A:Molecule type: DNA
 A:Residues: 792-6839 <BEN2>
 A:Cross-references: EMBL:X15423; NID:G6897; PID:CAA33463.1; PID:G6898
 A:Experimental source: var. Bristol
 R:Benian, G.M.; Kille, J.E.; Neckelmann, N.; Moertman, D.G.; Waterston, R.H.
 Nature 342, 45-50, 1989
 A:Title: Sequence of an unusually large protein implicated in regulation of myosin activity
 A:Reference number: S06797; MUID:90044042; PMID:2812002
 A:Accession: S06797
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 806-1175; 1178-1998; 'Y', 2000-3040; 'I', 3042-3335; 'I', 3337-5693; 5696-6359; 'I', 6361-6839
 A:Cross-references: EMBL:X15423
 A:Experimental source: var. Bristol
 R:Benian, G.M.; L'Hernault, S.W.; Morris, M.E.
 Genetics 134, 1097-1104, 1993
 A:Title: Additional sequence complexity in the muscle gene, unc-22, and its encoded protein
 A:Reference number: S57218; MUID:93387664; PMID:8397135
 A:Accession: S57218
 A:Molecule type: DNA
 A:Residues: 2-99; 108-194; 'Q', 196-206; 374-468; 658-753 <BEN3>
 A:Experimental source: var. Bristol
 R:White, S.
 Submitted to the EMBL Data Library, May 1996
 A:Reference number: Z20442
 A:Accession: T27934
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'MGIPGKCKGQ', 19-6839 <WIL>
 A:Cross-references: EMBL:Z73897; PIDN:CAA98064.1; GSPDB:GN00022; CESP:ZK617.1a
 A:Experimental source: clone ZK617
 R:Harlie, B.
 Submitted to the EMBL Data Library, May 1996
 A:Reference number: Z20458
 A:Accession: T28030
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 'MGIPGKCKGQ', 19-6839 <WIL>
 A:Cross-references: EMBL:Z73899; PIDN:CAA98081.1; GSPDB:GN00022; CESP:ZK617.1a
 A:Experimental source: clone ZK829
 C:Comment: Lack of unc-22 leads to a constant twitching of the body muscles.
 C:Genetics:
 A:Gene: unc-22; CESP:ZK617.1a
 A:Map position: 4
 A:Insertions: 18/3; 67/3; 143/2; 176/3; 264/2; 387/3; 413/2; 471/1; 516/3; 550/3; 582/3; 601/3; 669/3; 677/6/1; 680/3
 C:Superfamily: twitchin; fibronectin type III repeat homology; immunoglobulin homology;
 C:Keywords: ATP; autophosphorylation; duplication; muscle; phosphotransferase; serine/threonine kinase
 F:806-6898, 899-990, 991-1083, 1084-1175, 1178-1273, 1474-1567, 1770-1864, 2066-2158, 2358-2450, 2451-2540, 2541-2633, 2636-2678, 2681-2733, 2736-2833, 2836-2933, 2936-3033, 3036-3133, 3136-3233, 3236-3333, 3336-3433, 3436-3533, 3536-3633, 3636-3733, 3736-3833, 3836-3933, 3936-4033, 4036-4133, 4136-4233, 4236-4333, 4336-4433, 4436-4533, 4536-4633, 4636-4733, 4736-4833, 4836-4933, 4936-5033, 5036-5133, 5136-5233, 5236-5333, 5336-5433, 5436-5533, 5536-5633, 5636-5733, 5736-5833, 5836-5933, 5936-6033, 6036-6133, 6136-6233, 6236-6333, 6336-6433, 6436-6533, 6536-6633, 6636-6733, 6736-6833, 6836-6933, 6936-7033, 7036-7133, 7136-7233, 7236-7333, 7336-7433, 7436-7533, 7536-7633, 7636-7733, 7736-7833, 7836-7933, 7936-8033, 8036-8133, 8136-8233, 8236-8333, 8336-8433, 8436-8533, 8536-8633, 8636-8733, 8736-8833, 8836-8933, 8936-9033, 9036-9133, 9136-9233, 9236-9333, 9336-9433, 9436-9533, 9536-9633, 9636-9733, 9736-9833, 9836-9933, 9936-10033, 10036-10133, 10136-10233, 10236-10333, 10336-10433, 10436-10533, 10536-10633, 10636-10733, 10736-10833, 10836-10933, 10936-11033, 11036-11133, 11136-11233, 11236-11333, 11336-11433, 11436-11533, 11536-11633, 11636-11733, 11736-11833, 11836-11933, 11936-12033, 12036-12133, 12136-12233, 12236-12333, 12336-12433, 12436-12533, 12536-12633, 12636-12733, 12736-12833, 12836-12933, 12936-13033, 13036-13133, 13136-13233, 13236-13333, 13336-13433, 13436-13533, 13536-13633, 13636-13733, 13736-13833, 13836-13933, 13936-14033, 14036-14133, 14136-14233, 14236-14333, 14336-14433, 14436-14533, 14536-14633, 14636-14733, 14736-14833, 14836-14933, 14936-15033, 15036-15133, 15136-15233, 15236-15333, 15336-15433, 15436-15533, 15536-15633, 15636-15733, 15736-15833, 15836-15933, 15936-16033, 16036-16133, 16136-16233, 16236-16333, 16336-16433, 16436-16533, 16536-16633, 16636-16733, 16736-16833, 16836-16933, 16936-17033, 17036-17133, 17136-17233, 17236-17333, 17336-17433, 17436-17533, 17536-17633, 17636-17733, 17736-17833, 17836-17933, 17936-18033, 18036-18133, 18136-18233, 18236-18333, 18336-18433, 18436-18533, 18536-18633, 18636-18733, 18736-18833, 18836-18933, 18936-19033, 19036-19133, 19136-19233, 19236-19333, 19336-19433, 19436-19533, 19536-19633, 19636-19733, 19736-19833, 19836-19933, 19936-20033, 20036-20133, 20136-20233, 20236-20333, 20336-20433, 20436-20533, 20536-20633, 20636-20733, 20736-20833, 20836-20933, 20936-21033, 21036-21133, 21136-21233, 21236-21333, 21336-21433, 21436-21533, 21536-21633, 21636-21733, 21736-21833, 21836-21933, 21936-22033, 22036-22133, 22136-22233, 22236-22333, 22336-22433, 22436-22533, 22536-22633, 22636-22733, 22736-22833, 22836-22933, 22936-23033, 23036-23133, 23136-23233, 23236-23333, 23336-23433, 23436-23533, 23536-23633, 23636-23733, 23736-23833, 23836-23933, 23936-24033, 24036-24133, 24136-24233, 24236-24333, 24336-24433, 24436-24533, 24536-24633, 24636-24733, 24736-24833, 24836-24933, 24936-25033, 25036-25133, 25136-25233, 25236-25333, 25336-25433, 25436-25533, 25536-25633, 25636-25733, 25736-25833, 25836-25933, 25936-26033, 26036-26133, 26136-26233, 26236-26333, 26336-26433, 26436-26533, 26536-26633, 26636-26733, 26736-26833, 26836-26933, 26936-27033, 27036-27133, 27136-27233, 27236-27333, 27336-27433, 27436-27533, 27536-27633, 27636-27733, 27736-27833, 27836-27933, 27936-28033, 28036-28133, 28136-28233, 28236-28333, 28336-28433, 28436-28533, 28536-28633, 28636-28733, 28736-28833, 28836-28933, 28936-29033, 29036-29133, 29136-29233, 29236-29333, 29336-29433, 29436-29533, 29536-29633, 29636-29733, 29736-29833, 29836-29933, 29936-30033, 30036-30133, 30136-30233, 30236-30333, 30336-30433, 30436-30533, 30536-30633, 30636-30733, 30736-30833, 30836-30933, 30936-31033, 31036-31133, 31136-31233, 31236-31333, 31336-31433, 31436-31533, 31536-31633, 31636-31733, 31736-31833, 31836-31933, 31936-32033, 32036-32133, 32136-32233, 32236-32333, 32336-32433, 32436-32533, 32536-32633, 32636-32733, 32736-32833, 32836-32933, 32936-33033, 33036-33133, 33136-33233, 33236-33333, 33336-33433, 33436-33533, 33536-33633, 33636-33733, 33736-33833, 33836-33933, 33936-34033, 34036-34133, 34136-34233, 34236-34333, 34336-34433, 34436-34533, 34536-34633, 34636-34733, 34736-34833, 34836-34933, 34936-35033, 35036-35133, 35136-35233, 35236-35333, 35336-35433, 35436-35533, 35536-35633, 35636-35733, 35736-35833, 35836-35933, 35936-36033, 36036-36133, 36136-36233, 36236-36333, 36336-36433, 36436-36533, 36536-36633, 36636-36733, 36736-36833, 36836-36933, 36936-37033, 37036-37133, 37136-37233, 37236-37333, 37336-37433, 37436-37533, 37536-37633, 37636-37733, 37736-37833, 37836-37933, 37936-38033, 38036-38133, 38136-38233, 38236-38333, 38336-38433, 38436-38533, 38536-38633, 38636-38733, 38736-38833, 38836-38933, 38936-39033, 39036-39133, 39136-39233, 39236-39333, 39336-39433, 39436-39533, 39536-39633, 39636-39733, 39736-39833, 39836-39933, 39936-40033, 40036-40133, 40136-40233, 40236-40333, 40336-40433, 40436-40533, 40536-40633, 40636-40733, 40736-40833, 40836-40933, 40936-41033, 41036-41133, 41136-41233, 41236-41333, 41336-41433, 41436-41533, 41536-41633, 41636-41733, 41736-41833, 41836-41933, 41936-42033, 42036-42133, 42136-42233, 42236-42333, 42336-42433, 42436-42533, 42536-42633, 42636-42733, 42736-42833, 42836-42933, 42936-43033, 43036-43133, 43136-43233, 43236-43333, 43336-43433, 43436-43533, 43536-43633, 43636-43733, 43736-43833, 43836-43933, 43936-44033, 44036-44133, 44136-44233, 44236-44333, 44336-44433, 44436-44533, 44536-44633, 44636-44733, 44736-44833, 44836-44933, 44936-45033, 45036-45133, 45136-45233, 45236-45333, 45336-45433, 45436-45533, 45536-45633, 45636-45733, 45736-45833, 45836-45933, 45936-46033, 46036-46133, 46136-46233, 46236-46333, 46336-46433, 46436-46533, 46536-46633, 46636-46733, 46736-46833, 46836-46933, 46936-47033, 47036-47133, 47136-47233, 47236-47333, 47336-47433, 47436-47533, 47536-47633, 47636-47733, 47736-47833, 47836-47933, 47936-48033, 48036-48133, 48136-48233, 48236-48333, 48336-48433, 48436-48533, 48536-48633, 48636-48733, 48736-48833, 48836-48933, 48936-49033, 49036-49133, 49136-49233, 49236-49333, 49336-49433, 49436-49533, 49536-49633, 49636-49733, 49736-49833, 49836-49933, 49936-50033, 50036-50133, 50136-50233, 50236-50333, 50336-50433, 50436-50533, 50536-50633, 50636-50733, 50736-50833, 50836-50933, 50936-51033, 51036-51133, 51136-51233, 51236-51333, 51336-51433, 51436-51533, 51536-51633, 51636-51733, 51736-51833, 51836-51933, 51936-52033, 52036-52133, 52136-52233, 52236-52333, 52336-52433, 52436-52533, 52536-52633, 52636-52733, 52736-52833, 52836-52933, 52936-53033, 53036-53133, 53136-53233, 53236-53333, 53336-53433, 53436-53533, 53536-53633, 53636-53733, 53736-53833, 53836-53933, 53936-54033, 54036-54133, 54136-54233, 54236-54333, 54336-54433, 54436-54533, 54536-54633, 54636-54733, 54736-54833, 54836-54933, 54936-55033, 55036-55133, 55136-55233, 55236-55333, 55336-55433, 55436-55533, 55536-55633, 55636-55733, 55736-55833, 55836-55933, 55936-56033, 56036-56133, 56136-56233, 56236-56333, 56336-56433, 56436-56533, 56536-56633, 56636-56733, 56736-56833, 56836-56933, 56936-57033, 57036-57133, 57136-57233, 57236-57333, 57336-57433, 57436-57533, 57536-57633, 57636-57733, 57736-57833, 57836-57933, 57936-58033, 58036-58133, 58136-58233, 58236-58333, 58336-58433, 58436-58533, 58536-58633, 58636-58733, 58736-58833, 58836-58933, 58936-59033, 59036-59133, 59136-59233, 59236-59333, 59336-59433, 59436-59533, 59536-59633, 59636-59733, 59736-59833, 59836-59933, 59936-60033, 60036-60133, 60136-60233, 60236-60333, 60336-60433, 60436-60533, 60536-60633, 60636-60733, 60736-60833, 60836-60933, 60936-61033, 61036-61133, 61136-61233, 61236-61333, 61336-61433, 61436-61533, 61536-61633, 61636-61733, 61736-61833, 61836-61933, 61936-62033, 62036-62133, 62136-62233, 62236-62333, 62336-62433, 62436-62533, 62536-62633, 62636-62733, 62736-62833, 62836-62933, 62936-63033, 63036-63133, 63136-63233, 63236-63333, 63336-63433, 63436-63533, 63536-63633, 63636-63733, 63736-63833, 63836-63933, 63936-64033, 64036-64133, 64136-64233, 64236-64333, 64336-64433, 64436-64533, 64536-64633, 64636-64733, 64736-64833, 64836-64933, 64936-65033, 65036-65133, 65136-65233, 65236-65333, 65336-65433, 65436-65533, 65536-65633, 65636-65733, 65736-65833, 65836-65933, 65936-66033, 66036-66133, 66136-66233, 66236-66333, 66336-66433, 66436-66533, 66536-66633, 66636-66733, 66736-66833, 66836-66933, 66936-67033, 67036-67133, 67136-67233, 67236-67333, 67336-67433, 67436-67533, 67536-67633, 67636-67733, 67736-67833, 67836-67933, 67936-68033, 68036-68133, 68136-68233, 68236-68333, 68336-68433, 68436-68533, 68536-68633, 68636-68733, 68736-68833, 68836-68933, 68936-69033, 69036-69133, 69136-69233, 69236-69333, 69336-69433, 69436-69533, 69536-69633, 69636-69733, 69736-69833, 69836-69933, 69936-70033, 70036-70133, 70136-70233, 70236-70333, 70336-70433, 70436-70533, 70536-70633, 70636-70733, 70736-70833, 70836-70933, 70936-71033, 71036-71133, 71136-71233, 71236-71333, 71336-71433, 71436-71533, 71536-71633, 71636-71733, 71736-71833, 71836-71933, 71936-72033, 72036-72133, 72136-72233, 72236-72333, 72336-72433, 72436-72533, 72536-72633, 72636-72733, 72736-72833, 72836-72933, 72936-73033, 73036-73133, 73136-73233, 73236-73333, 73336-73433, 73436-73533, 73536-73633, 73636-73733, 73736-73833, 73836-73933, 73936-74033, 74036-74133, 74136-74233, 74236-74333, 74336-74433, 74436-74533, 74536-74633, 74636-74733, 74736-74833, 74836-74933, 74936-75033, 75036-75133, 75136-75233, 75236-75333, 75336-75433, 75436-75533, 75536-75633, 75636-75733, 75736-75833, 75836-75933, 75936-76033, 76036-76133, 76136-76233, 76236-76333, 76336-76433, 76436-76533, 76536-76633, 76636-76733, 76736-76833, 76836-76933, 76936-77033, 77036-77133, 77136-77233, 77236-77333, 77336-77433, 77436-77533, 77536-77633, 77636-77733, 77736-77833, 77836-77933, 77936-78033, 78036-78133, 78136-78233, 78236-78333, 78336-78433, 78436-78533, 78536-78633, 78636-78733, 78736-78833, 78836-78933, 78936-79033, 79036-79133, 79136-79233, 79236-79333, 79336-79433, 79436-79533, 79536-79633, 79636-79733, 79736-79833, 79836-79933, 79936-80033, 80036-80133, 80136-80233, 80236-80333, 80336-80433, 80436-80533, 80536-80633, 80636-80733, 80736-80833, 80836-80933, 80936-81033, 81036-81133, 81136-81233, 81236-81333, 81336-81433, 81436-81533, 81536-81633, 81636-81733, 81736-81833, 81836-81933, 81936-82033, 82036-82133, 82136-82233, 82236-82333, 82336-82433, 82436-82533, 82536-82633, 82636-82733, 82736-82833, 82836-82933, 82936-83033, 83036-83133, 83136-83233, 83236-83333, 83336-83433, 83436-83533, 83536-83633, 83636-83733, 83736-83833, 83836-83933, 83936-84033, 84036-84133, 84136-84233, 84236-84333, 84336-84433, 84436-84533, 84536-84633, 84636-84733, 84736-84833,

```
Db      1465 IGIENKNEADVVEIDGAGLRGSCNVVLEAEKRPILNMKKKILEATAG-EPCVVXVPFOI 1523
      211 TGSALDPDQ 219
      1524 KGTTRGDDPK 1532

RESULT 28
503199
opioid-binding protein OPCAM precursor - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 17-Mar-2000
C:Accession: S03199
R:Schotfield, P.R.; McFarland, K.C.; Haylick, J.S.; Wilcox, J.N.; Cho, T.M.; Roy, S.; Le
EMBO J. 8, 489-495, 1989
A:Title: Molecular characterization of a new immunoglobulin superfamily protein with pot
A:Reference number: S03199; MUID:9251576; PMID:2721489
A:Accession: S03199
A:Molecule type: mRNA
A:Residues: 1345 <SCH>
A:Cross-references: EMBL:X12672; NID:9585; PIDN:CAA31192.1; PID:9586
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termin
C:Keywords: transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-345/Product: opioid-binding protein OPCAM #status predicted <MAT>

Query Match      8.9%; Score 117.5; DB 2; Length 345;
Best Local Similarity 22.4%; Pred. No. 0.25;
Matches 59; Conservative 43; Mismatches 100; Indels 61; Gaps 10;

QY      11 LVVLQALLPAAQGNKVVG-----KKGDTELTCTASQKSIQFMKNSN 57
      14 LVVSLRLFLVPTGVPRSGDATPPAMDNVTYQGESATLRTCTDDRT-RVAMLNRS 72
      58 QIKILGN-----QGSFLTKGPSKLNDRASRRSLWDQGNFLLIKLIEDSDTYICE 110
      73 TLVAGNDKMSIDPRVILVNTPTQ-----YSMIQNVVDYEGPYTCS 116
QY      111 VE----DQKEEVQLLVFG---LTNSDTHLLQGOSLTTLTLESPGSSPSVQCRSPRGKN 162
      117 VQTDNHEKTSRVHLIVGVPPQIMNISDVTVNEGSSVTLTCLAIGRPEPTVTRHLSVKE 176
QY      163 IQG---GKTLVSQLELDQSGTWCTVLO---NOKVYEFKIDYPRASALPAPPTGS 213
      177 GQGFVSDEYLEISDTRDQSGEYECALNDVAAPDVKVKITVNPYIS--KAKNTGV 234
Db      214 ALPDPQTASALPDPAPASALPAA 236
      235 SVQGGKILSC-----EASAVPMA 252

RESULT 29
A41288
vascular cell adhesion molecule 1, long splice form precursor - human
N:Alternate names: VCAM-1
C:Species: Homo sapiens (man)
C>Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 31-Jan-2000
C:Accession: A41288; S11476; A39755; B39755; A61160; A43352; PH1379; A39554
R:Cybulsky, M.I.; Fries, J.W.U.; Williams, A.J.; Sultan, P.; Eddy, R.; Byers, M.; Shows,
Proc. Natl. Acad. Sci. U.S.A. 88, 7859-7863, 1991
A:Title: Gene structure, chromosomal location, and basis for alternative mRNA splicing c
A:Reference number: A41288; MUID:91352090; PMID:11715583
A:Accession: A41288
A:Molecule type: DNA
A:Residues: 1-739 <CYB>
A:Cross-references: GB:M73255; NID:9340195; PIDN:AAA61270.1; PID:9340196
R:Polte, T.; Newman, W.; Gopal, T.V.
Nucleic Acids Res. 18, 5901, 1990
A:Title: Full length vascular cell adhesion molecule 1 (VCAM-1).
A:Reference number: S11476; MUID:91016951; PMID:1699207
A:Accession: S11476
A:Status: translation not shown

A:Molecule type: mRNA
A:Residues: 1-739 <BOU>
A:Cross-references: EMBL:X53051; NID:937648; PIDN:CAA37218.1; PID:937649
R:Hession, C.; Tizard, R.; Vassallo, C.; Schiffer, S.B.; Goff, D.; Moy, P.; Chi-Rosso,
J. Biol. Chem. 266, 6682-6685, 1991
A:Title: Cloning of an alternate form of vascular cell adhesion molecule-1 (VCAM1).
A:Reference number: A39755; MUID:91201302; PMID:1107873
A:Accession: A39755
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 25-646, 648-739 <HES1>
A:Cross-references: GB:M60335
A:Note: the complete translation is not shown
A:Accession: B39755
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-739 <HES2>
A:Cross-references: GB:M60335; NID:9340193; PIDN:AAA61269.1; PID:9340194
A:Experimental source: cell type endothelial cell; tissue type umbilical vein; map 1p32
R:Cybulsky, M.I.; Fries, J.W.U.; Williams, A.J.; Sultan, P.; Davis, V.M.; Gimbrone Jr.,
Am. J. Pathol. 138, 815-820, 1992
A:Title: Rapid communication. Alternative splicing of human VCAM-1 in activated vascul
A:Reference number: A61160; MUID:91189297; PMID:1107234
A:Accession: A61160
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 25-401, 'T', 403-686 <CY2>
R:Adamarco, M.F.; McQuillan, J.O.; Rosen, G.D.; Dean, D.C.
J. Biol. Chem. 267, 16323-16329, 1992
A:Title: Characterization of the promoter for vascular cell adhesion molecule-1 (VCAM-1)
A:Reference number: A43352; MUID:92355594; PMID:1179595
A:Accession: A43352
A:Molecule type: DNA
A:Residues: 1-21 <IND>
A:Cross-references: GB:M92431; NID:9340197
A:Note: sequence extracted from NCBI backbone (NCBIN:110680, NCBI:P.110681)
R:Oborn, L.; Vassallo, C.; Benjamin, C.D.
J. Exp. Med. 176, 99-107, 1992
A:Title: Activated endothelium binds lymphocytes through a novel binding site in the al
A:Reference number: PH1379; MUID:92308860; PMID:1377228
A:Accession: PH1379
A:Molecule type: protein
A:Residues: 25-181, 'G', 183-402 <OSB>
C:Comment: This adhesion molecule is induced on endothelial cells by inflammatory cytok
C:Genetics:
A:Gene: GDB:VCAM1
A:Cross-references: GDB:127922; OMIM:192225
A:Map position: 1p32-1p31
C:Keywords: alternative splicing; cell adhesion; glycoprotein; transmembrane protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-739/Product: vascular cell adhesion molecule 1, long splice form #status predicted
F:25-698/Domain: extracellular #status predicted <EXT>
F:699-720/Domain: transmembrane #status predicted <TM>
F:721-739/Domain: intracellular #status predicted <INT>
F:273,365,417,463,531,561/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      8.9%; Score 117; DB 2; Length 739;
Best Local Similarity 22.3%; Pred. No. 0.67;
Matches 47; Conservative 38; Mismatches 84; Indels 42; Gaps 7;

QY      8 RHLLVQLALPAA--GKNVVLGKGDVETLTCTASQKSIQFMKNSNQIKILGNQ 65
      211 RQAVKRLQVYISKNTVIVNPTKQBGGSVMTCSGSLPAPLFW----- 258
      66 GSFVTGPSKLNDRASRRSLWDQGNFLLIKLIEDSDTYICE---VEDQKEVQL 121
      259 -----SKLDNGNLMHL---SGNATLTLLAMRMEDSGIYCEGVNLTGKRXKEVELI 307
QY      122 V----FGTANSTHLLQ--GOSLTTLTLESPGSSPSVQCRSPRGKNIOG-----GKT 168
      308 VQKPTVEISPPRIAQIGDSVMLTCSVMGCESSFSFWRTQIDSPLSGVKVSGETNST 367
QY      169 LSVSQLELDQSGTWCTVLOQNKVEFKIDI 199
```

Db 368 LTLSPVSEFNEHSYLCVTGCHKKLEKGIQV 398

RESULT 30

113669 neuromusculin - fruit fly (*Drosophila melanogaster*)

C/Species: *Drosophila melanogaster*

C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C/Accession: J13669

R/Kania, A.; Han, P.L.; Kim, Y.T.; Bellen, H.

Neuron 11, 673-687, 1993

A/Title: Neuromusculin, a *Drosophila* gene expressed in peripheral neuronal precursors at

A/Reference number: Z17697; MUID:9400831; PMID:8398154

A/Accession: J13669

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1011 <KAN>

A/Cross-references: EMBL:J23146; NID:G385073; PID:G385074; PIDN:AAA03750.1

C/Genetics:

A/Gene: nrm

A/Cross-references: FlyBase:FBgn0005629

Query Match 8.9%; Score 117; DB 2; Length 1011;

Best Local Similarity 25.0%; Pred. No. 0.97;

Matches 48; Conservative 38; Mismatches 76; Indels 30; Gaps 9;

Qy 10 LLLVQLALPAATQGNKRVLLGKGGDTVELCTASQKSIQFH---WKNSTNQT---KIL 62

Db 26 LVLVCLALVDSSTQAVDTTISQSGSGVLPCEPDAKCGCKLSLWPFKGGDDRIAMLL 85

Qy 63 GNQGSFLTKGPKSKNDRAISRSLMDQGNFPLIKNKIEDSDTYICEVEDQKEVQLLV 122

Db 86 GD-----SNVTSVAKKEDERVTV-EQNPYRLVIXDKLAEEDLYLCTP-----T 128

Qy 123 FGLTANSDTHLQGSLLTLTLESPGSSPSVQCSPPGSKNIQGGKTLVSQLELDQSGTW 182

Db 129 FPIEBETCDN-FNGRIELRLVLP---TEVVIDAKDKIRKNSV--VGMQGRQSLKA 182

Qy 183 TCTVLQNGKKVE 194

Db 183 TCTVNRTRPQPE 194

RESULT 31

JC4025 oploid-binding cell adhesion protein - human

C/Species: *Homo sapiens* (man)

C/Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 17-Mar-2000

C/Accession: J04025

R/Sharp, K.B.; Lee, N.M.

Gene 155, 213-217, 1995

A/Title: Cloning, sequencing and localization to chromosome 11 of a cDNA encoding a huma

A/Reference number: J04025; MUID:95237612; PMID:7721093

A/Accession: J04025

A/Molecule type: mRNA

A/Residues: 1-345 <SHA>

A/Cross-references: GB:J24774; NID:G514373; PIDN:AAA36387.1; PID:G514374

A/Experimental source: brain

C/Comment: This protein binds oploid alkaloids in the presence of acidic lipids, exhibit

C/Genetics:

A/Gene: GDB:OPCML; OBCAM; OPCM

A/Cross-references: GDB:251677; OMIM:600632

A/Map position: 11pter-11qter

C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi

Query Match 8.8%; Score 116.5; DB 2; Length 345;

Best Local Similarity 22.4%; Pred. No. 0.3; Mismatches 100; Indels 61; Gaps 10;

Matches 59; Conservative 43; Mismatches 100; Indels 61; Gaps 10;

Qy 11 LLLVQLALPAATQGNKRVLLG-----KKGDTVELCTASQKSIQFHWKSN 57

Db 14 LVVSLRLFLVPGVPRSGDAPFPKAMVNTVROGSSATLRCTIDRVT-RVAMNRS 72

Qy 58 QIKILGN-----QGSFLTKGPKSKLNDRAISRSLMDQGNFPLIKNKIEDSDTYICE 110

Db 73 TILVAGNDKSIDRVLIVNTPTQ-----YSIMIQVNDVYDGPYTC 116

Qy 111 VE-----DQKEVQLLVFG---LTRANSPTHLLQGSLLTLTLESPGSSPSVQCSPPGSKN 162

Db 117 VQTNHRTKTSVHLIVQPPQIMMISDITWESSVTLLCLAGRPETVTWHLVSKE 176

Qy 163 IQG---GKTLVSQLELDQSGTWCTVLO-----NOKVFEKIDIVPRASALPAPPGS 213

Db 177 GQGVSSDEYLEISDIRDQSGEYECALNDVAAPDVAKVITVNPYPYS--KAKVTGV 234

Qy 214 ALPPQFASALPDPAPASALPAA 236

Db 235 SVGGKGLISC---EASAVPMA 252

RESULT 32

B41288 vascular cell adhesion molecule 1, short splice form precursor - human

N/Alternate names: VCAM-1

C/Species: *Homo sapiens* (man)

C/Date: 03-Apr-1992 #sequence_revision 14-Jul-1994 #text_change 31-Jan-2000

C/Accession: B41288; A33758

R/Cybulsky, M.I.; Fries, J.W.U.; Williams, A.J.; Sultan, P.; Eddy, R.; Byers, M.; Shows

Proc. Natl. Acad. Sci. U.S.A. 88, 7859-7863, 1991

A/Title: Gene structure, chromosomal location, and basis for alternative mRNA splicing

A/Reference number: A41288; MUID:91352090; PMID:1715583

A/Accession: B41288

A/Molecule type: DNA

A/Residues: 1-647 <CYB>

A/Cross-references: GB:M73255

R/Osborn, L.; Hession, C.; Tizard, R.; Vassallo, C.; Lubowsky, S.; Chi-Rosso, G.; Lobb

Cell 59, 1203-1211, 1989

A/Title: Direct expression cloning of vascular cell adhesion molecule 1, a cytokine-ind

A/Reference number: A33758; MUID:9090619; PMID:2688898

A/Accession: A33758

A/Molecule type: mRNA

A/Residues: 1-647 <OSB>

A/Cross-references: GB:M30257; NID:G179885; PIDN:AAA51917.1; PID:G179886

A/Keywords: alternative splicing; cell adhesion; glycoprotein; transmembrane protein

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-647/Product: vascular cell adhesion molecule 1, short form #status predicted <Mnt>

F:25-606/Domain: extracellular #status predicted <EXT>

F:607-628/Domain: transmembrane #status predicted <TM>

F:629-647/Domain: intracellular #status predicted <INT>

F:273,325,371,439,469/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 8.7%; Score 115; DB 2; Length 647;

Best Local Similarity 20.3%; Pred. No. 0.8;

Matches 61; Conservative 49; Mismatches 75; Indels 116; Gaps 13;

Qy 16 LALLPAAATQGNKV-----VLGGKDTVELCTASQKSIQFHWKSNQIKIG---NQ 65

Db 15 LMIIPASQAQKIEETTESKRLAQIGDSVLSITCSPTCCSPFBSKRIQDISPLNGKXTNE 74

Qy 66 G-SPLTKGPKSKLNDRAISRSLMDQGNFPLIKNKIEDSDTYICEVED--- 93

Db 75 GTTSTLTMMNPFVSEFNEHSYLCATNCESRK--LEKGIQVEIYSPKDEIHLGPLEAGKP 132

Qy 94 -----LTIKULK-----IEDSD-----TYICEVED--- 113

Db 133 ITVKGSVADYPPEDRLIEDLKGDLHMKSGQFEDADARKLETSLVETFTPVLEDIGKV 192

Qy 114 -----QKEVQLLVFG---LTRANSPTHLLQGSLLTLTLESPGSS 149

Db 193 LVCAKALHIDBMSVPTVRQAVKELQYISPKNTIVISVNSTKLQEGSSVTMTCSSEGLP 252

Qy 150 SPSPV---QCSPPGSKNIQGGKTLVSQLELDQSGTWCT---TVLQNGKVEFKIDIVP 201

Db 253 APEIFMGKLDNGMLQHLGSLGATLTLLAMMEMDSGIVCGVNLIGNRRKEVELIVQAF 312

Db 58 APKLLIYGASNLSEGVPSRFGSGGCTDFTLTITSLQGDPSATYCCQ 104

RESULT 36

A:Accession: A6319

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-37 <RES>

A:Cross-references: GB:M60964; NID:g180215; PIDN:AAA51964.1; PID:g180217

R:Zimmermann, W.; Ortlieb, B.; Friedlrich, R.; von Kleist, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 2960-2964, 1987

A:Title: Isolation and characterization of cDNA clones encoding the human carcinoembryo-

A:Reference number: 159098; MUID:87204247; PMID:3033671

A:Accession: 159098

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 331-702 <RES>

A:Cross-references: GB:M6234; NID:g180240; PIDN:AAA51972.1; PID:g180241

R:Stephen, D.; Paxton, R.J.; Neumaier, M.; Shively, J.E.; Wagener, C.

Biochem. Biophys. Res. Commun. 147, 212-218, 1987

A:Title: Carcinoembryonic antigen (CEA) and two crossreacting antigens of 165 KD and 10

A:Reference number: A26831; MUID:87326349; PMID:3632664

A:Accession: A26831

A:Molecule type: protein

A:Residues: 35-64 <SIE>

R:Thomas, P.; Toth, C.A.

Biochem. Biophys. Res. Commun. 170, 391-396, 1990

A:Title: Carcinoembryonic antigen binding to Kupffer cells is via a peptide located at

A:Reference number: A35490; MUID:90321257; PMID:2372227

A:Accession: A35490

A:Molecule type: protein

A:Residues: 1-702 <BEA>

A:Cross-references: GB:M29540; NID:g180222; PIDN:AAA51967.1; PID:g180223

R:Barnett, T.; Goeddel, S.J.; Nothdurft, M.A.; Elting, J.J.

Genomic 3, 59-66, 1988

A:Title: Carcinoembryonic antigen family: characterization of cDNAs coding for NCA and C

A:Reference number: A31037; MUID:89122014; PMID:3220478

A:Accession: A31037

A:Molecule type: mRNA

A:Residues: 1-702 <BAR>

A:Cross-references: GB:M29540; NID:g180222; PIDN:AAA51967.1; PID:g180223

A:Note: the authors translated the codon GTG for residue 130 as Leu

R:Okawa, S.; Nakazato, H.; Kosaki, G.

Biochem. Biophys. Res. Commun. 142, 511-518, 1987

A:Title: Primary structure of human carcinoembryonic antigen (CEA) deduced from cDNA seq

A:Reference number: A25845; MUID:87128144; PMID:3814146

A:Accession: A25845

A:Molecule type: mRNA

A:Residues: 5-702 <OIK>

A:Cross-references: GB:M15042; NID:g180198; PIDN:AAA51963.1; PID:g180199

R:Okawa, S.

Submitted to the EMBL Data Library, September 1989

A:Reference number: S08106

A:Accession: S08106

A:Molecule type: mRNA

A:Residues: 5-319, 321-702 <OIZ>

A:Cross-references: EMBL:X16455; NID:g29854; PIDN:CAA34474.1; PID:g825638

R:Barnett, T.

Submitted to the EMBL Data Library, September 1991

A:Description: Genomic DNA sequence upstream of the translational start of the carcinoemb

A:Reference number: S31737

A:Accession: S31737

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-141 <BAZ>

A:Cross-references: EMBL:X62151

R:Khan, W.N.; Prangemeyer, L.; Teglund, S.; Israelson, A.; Bremer, K.; Hammarstrom, S.

Genomic 14, 394-390, 1992

A:Title: Identification of three new genes and estimation of the size of the carcinoemb

A:Reference number: A44476; MUID:93052339; PMID:1427854

A:Accession: A44476

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 35-141 <KHA>

R:Wilcocke, T.C.; Craig, I.W.

Genomic 8, 492-500, 1990

A:Title: Characterization of the genomic organization of human carcinoembryonic antigen

A:Reference number: 154224; MUID:91139118; PMID:2266372

A:Accession: 154224

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-37 <RES>

A:Cross-references: GB:M60964; NID:g180215; PIDN:AAA51964.1; PID:g180217

R:Zimmermann, W.; Ortlieb, B.; Friedlrich, R.; von Kleist, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 2960-2964, 1987

A:Title: Isolation and characterization of cDNA clones encoding the human carcinoembryo-

A:Reference number: 159098; MUID:87204247; PMID:3033671

A:Accession: 159098

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 331-702 <RES>

A:Cross-references: GB:M6234; NID:g180240; PIDN:AAA51972.1; PID:g180241

R:Stephen, D.; Paxton, R.J.; Neumaier, M.; Shively, J.E.; Wagener, C.

Biochem. Biophys. Res. Commun. 147, 212-218, 1987

A:Title: Carcinoembryonic antigen (CEA) and two crossreacting antigens of 165 KD and 10

A:Reference number: A26831; MUID:87326349; PMID:3632664

A:Accession: A26831

A:Molecule type: protein

A:Residues: 35-64 <SIE>

R:Thomas, P.; Toth, C.A.

Biochem. Biophys. Res. Commun. 170, 391-396, 1990

A:Title: Carcinoembryonic antigen binding to Kupffer cells is via a peptide located at

A:Reference number: A35490; MUID:90321257; PMID:2372227

A:Accession: A35490

A:Molecule type: protein

A:Residues: 1-702 <BEA>

A:Cross-references: GB:M29540; NID:g180222; PIDN:AAA51967.1; PID:g180223

R:Barnett, T.; Goeddel, S.J.; Nothdurft, M.A.; Elting, J.J.

Genomic 3, 59-66, 1988

A:Title: Carcinoembryonic antigen family: characterization of cDNAs coding for NCA and C

A:Reference number: A31037; MUID:89122014; PMID:3220478

A:Accession: A31037

A:Molecule type: mRNA

A:Residues: 1-702 <BAR>

A:Cross-references: GB:M29540; NID:g180222; PIDN:AAA51967.1; PID:g180223

A:Note: the authors translated the codon GTG for residue 130 as Leu

R:Okawa, S.; Nakazato, H.; Kosaki, G.

Biochem. Biophys. Res. Commun. 142, 511-518, 1987

A:Title: Primary structure of human carcinoembryonic antigen (CEA) deduced from cDNA seq

A:Reference number: A25845; MUID:87128144; PMID:3814146

A:Accession: A25845

A:Molecule type: mRNA

A:Residues: 5-702 <OIK>

A:Cross-references: GB:M15042; NID:g180198; PIDN:AAA51963.1; PID:g180199

R:Okawa, S.

Submitted to the EMBL Data Library, September 1989

A:Reference number: S08106

A:Accession: S08106

A:Molecule type: mRNA

A:Residues: 5-319, 321-702 <OIZ>

A:Cross-references: EMBL:X16455; NID:g29854; PIDN:CAA34474.1; PID:g825638

R:Barnett, T.

Submitted to the EMBL Data Library, September 1991

A:Description: Genomic DNA sequence upstream of the translational start of the carcinoemb

A:Reference number: S31737

A:Accession: S31737

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-141 <BAZ>

A:Cross-references: EMBL:X62151

R:Khan, W.N.; Prangemeyer, L.; Teglund, S.; Israelson, A.; Bremer, K.; Hammarstrom, S.

Genomic 14, 394-390, 1992

A:Title: Identification of three new genes and estimation of the size of the carcinoemb

A:Reference number: A44476; MUID:93052339; PMID:1427854

A:Accession: A44476

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 35-141 <KHA>

R:Wilcocke, T.C.; Craig, I.W.

Genomic 8, 492-500, 1990

A:Title: Characterization of the genomic organization of human carcinoembryonic antigen

A:Reference number: 154224; MUID:91139118; PMID:2266372

Query Match 8.44; Score 111; DB 2; Length 702;

Best Local Similarity 20.96; Pred. No. 1.7;

Matches 51; Conservative 41; Mismatches 98; Indels 54; Gaps 9;

20 PAATGGKVVILGKGGDTVELCTASOKKSIQFMKNSNOIKITGNGSFLTKGSKLND 79

503 PSISNSKRYEDK-DVATCEBEAQTYYLWVNGQS.LV----- 543

80 ADSRRSLMDQNFPLIKLKIEDSDTYICEV-----DQKEEYOL-LVFG---LTANS 129

544 --SPRLDLSNGNRTLTLEFVTRNDARAYVCGIANSVANSNDPVLTLVGPDPRLSP 601

130 DTHLQGSQSLTTLSPSSSPVQCSPPGKXIKIGSKITLSVQLBDSGTWTCTYLQN 189

602 DSYLSGANLNLCHSANSNPYSWMT-NGIQOHQYVFIKIPNNNGTYACFY--- 657

190 QKVEFIDIVPPASALPAPPTGALPDPPTASALPAPASALPALAVISFLGLG 249

658 -----SNL-ATGRNNSIVKSLTVASAGSPSL---AGATVGMIGVLVG 698

250 VACV 253

699 VALI 702

RESULT 37

S33850
fibronectin-binding protein - Streptococcus dysgalactiae

C:Species: Streptococcus dysgalactiae

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 15-Oct-1999

C:Accession: S33850, S33632

R:Lindgren, P.E.; McGavin, M.J.; Signaes, C.; Guss, B.; Gurnaidappa, S.; Hoocek, M.; Li

Eur. J. Biochem. 214, 819-827, 1993

A:Title: Two different genes coding for fibronectin-binding proteins from Streptococcus

A:Reference number: S33850, MUID:93307299, PMID:8319691

A:Accession: S33850

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1091 <LIN>

A:Cross-references: EMBL:Z22150, NID:9288968, PIDN:CAA80121.1, PID:9288969

Query Match

Best Local Similarity 23.0%; Score 111; DB 2; Length 1091;

Matches 64; Conservative 41; Mismatches 107; Indels 66; Gaps 11;

34 GDTVELTCTASQKKSIGFHMKNNOIKIL---GNQSSFLTKGFSKLNDRSRSLMDQ 89

821 GQSEIRITEDSGQMGSGQPGSGNETVVEDTQSGEDIVLGGPGQVIDFTEDSQPMMSG 880

90 GNPFLLIKNLKIEDSDTYIC---EVEDQKEVQLVFGLTANSDTHL----- 134

881 NNSHTIEDSKPSQEDSVIIGGQGVVIDFTEDTQSGSGNSHDTGTVLEEDSKPSQED 940

135 ---OCQSLTLTLESPGSSPSVQCRSP-----RGKNIQGGKTLVSQLELD--- 178

941 VIIGGQGVVIDFTEDTQSGAGQVSPITTEBTHKPEILMGQSPIMWE--DTLPG 998

179 -SGTWTCTVLQ--NQKVEFKIDIVPRASALPAP-PTGSALPPQRTASALPDPASALP 234

999 MSGSENVATVEEDRPRLQFHF--NEEPVATVPTVSGTPIAQVSKVPHAKASALP 1055

235 -----AALAVTSFLGLGLGVAACVIAKRR 258

1056 QTGDTNKLTFETFTTALTTLV-----GAAGLGGKR 1085

Db

RESULT 38

149443
gene 284 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999

C:Accession: 149443

R:Mathew, P.A.; Garni-Wagner, B.A.; Land, K.; Takashima, A.; Stoneman, E.; Bennett, M.;

J. Immunol. 151, 5328-5337, 1993

A:Title: Cloning and characterization of the 284 gene encoding a molecule associated with

A:Reference number: 149443, MUID:94044757, PMID:8228228

A:Accession: 149443

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-398 <RES>

A:Cross-references: GB:L19057, NID:9309076, PIDN:AAA16353.1, PID:9309077

A:Genetics:

A:Gene: 284

Query Match

Best Local Similarity 22.7%; Score 110.5; DB 2; Length 398;

Matches 51; Conservative 38; Mismatches 87; Indels 49; Gaps 9;

18 LLPATQ-----NKVLGKGGDTVELTCTASQKKSIGFHMKNNO-----IKILNQG 66

12 LILRAHQGQCPDSEEVVGVSKPQLRPSNIGTKOVSVQMKTEGSHRKIEIL----- 67

67 SFLTKGSKLNDRSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEV-----QLL 121

68 NMYNDGFSWNSVPSDIYG--FDYDGFALSTKSATLQDSGHYLLTITWTGVCVKNFQL 126

122 VFGLTANSPTHLQGGSLTLTLESPPSSPSVQ-----CRSRGNKI----- 163

Db

Query Match

Best Local Similarity 23.0%; Score 110.5; DB 1; Length 2029;

Matches 50; Conservative 37; Mismatches 75; Indels 55; Gaps 11;

20 PATQGNKVLGKGGDTVELTCTASQKKSIGFHMKNNOIKILNGGSLTKGFSKLNDR 79

140 PVTQPGRGVLEVHTVMTCAINPNTIY--IKQTKV-----DM 182

80 ADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDTHL--LOGS 138

183 SNRYSLKD--GR-LQIENSREDQGYECVAENS-----MGREHSKATNLVYKVR 231

139 LITLESPP-----GSPSVQCRS-----PRGNIGGKTL-----VSQ 174

232 VPPTFSRPETISEVMGLGNLNLCTAVSGPMPHVMMKGSBDLTENEMPIGRNVLQ 291

175 ELQDSGTWTCTVLQNGKVEFKIDIVPRASALPAPT 211

Db

127 IL-----DHY---ETPLKQKQKRWINGTQQLPLSLCVLTADNDVSYAFWRGSLTISN 176

164 QGKTLVSQLELDQSGTWTCTVLQNGKVEFKIDIVPRASALPA 208

177 QNRSTHMQIDASSLHTYTICNVSRASWANTLTNTHCCQSVPS 221

Db

RESULT 39

TDFFLK

protein-tyrosine-phosphatase (EC 3.1.3.48) DLR precursor - fruit fly (Drosophila melan-

N:Alternate names: leukocyte antigen-related protein

C:Species: Drosophila melanogaster

C:Date: 14-Dec-1990 #sequence_revision 02-May-1994 #text_change 22-Jun-1999

C:Accession: A36182

R:Streuli, M.; Krueger, N.X.; Tsai, A.Y.M.; Saito, H.

Proc. Natl. Acad. Sci. U.S.A. 86, 8698-8702, 1989

A:Title: A family of receptor-linked protein tyrosine phosphatases in humans and Drosopi-

A:Reference number: A36182, MUID:90046860, PMID:2554325

A:Accession: A36182

A:Molecule type: mRNA

A:Residues: 1-2029 <STR>

A:Cross-references: GB:M27700, NID:9157811, PIDN:AAA28668.1, PID:9157812

C:Genetics:

A:Gene: FlyBase:lar

A:Cross-references: FlyBase:FBgn0000464

C:Superfamily: Leukocyte antigen-related protein; fibronectin type III repeat homology;

ogy

C:Keywords: glycoprotein; phosphoprotein; phosphoric monoester hydrolase; transmembrane

F:1-32/Domain: signal sequence #status predicted <SIG>

F:33-2029/Product: leukocyte antigen-related protein #status predicted <MAT>

F:33-1377/Domain: extracellular #status predicted <EXT>

F:50-113/Domain: immunoglobulin homology <IMM1>

F:154-211/Domain: immunoglobulin homology <IMM2>

F:249-303/Domain: fibronectin type III repeat homology <FN3>

F:321-401/Domain: fibronectin type III repeat homology <FN3>

F:416-502/Domain: fibronectin type III repeat homology <FN3>

F:514-599/Domain: fibronectin type III repeat homology <FN3>

F:610-699/Domain: fibronectin type III repeat homology <FN3>

F:708-802/Domain: fibronectin type III repeat homology <FN3>

F:811-896/Domain: fibronectin type III repeat homology <FN3>

F:909-993/Domain: fibronectin type III repeat homology <FN3>

F:1006-1091/Domain: fibronectin type III repeat homology <FN3>

F:1101-1198/Domain: fibronectin type III repeat homology <FN3>

F:1378-1402/Domain: transmembrane #status predicted <TM>

F:1417-2029/Domain: intracellular #status predicted <INT>

F:1417-2029/Domain: leukocyte common antigen cytosolic domain homology <LAC>

F:1786-2009/Domain: protein-tyrosine-phosphatase homology <PTP2>

F:57-111,161-209,256-301/Dissulfide bonds: #status predicted

F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:1670/Active site: Cys (phosphocysteine intermediate) #status predicted

F:1676/Binding site: substrate phosphate (Arg) #status predicted

F:1661/Active site: Cys (phosphocysteine intermediate) #status predicted

F:1967/Binding site: substrate phosphate (Arg) #status predicted

Query Match

Best Local Similarity 23.0%; Score 110.5; DB 1; Length 2029;

Matches 50; Conservative 37; Mismatches 75; Indels 55; Gaps 11;

20 PATQGNKVLGKGGDTVELTCTASQKKSIGFHMKNNOIKILNGGSLTKGFSKLNDR 79

140 PVTQPGRGVLEVHTVMTCAINPNTIY--IKQTKV-----DM 182

80 ADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEVQLVFGLTANSDTHL--LOGS 138

183 SNRYSLKD--GR-LQIENSREDQGYECVAENS-----MGREHSKATNLVYKVR 231

139 LITLESPP-----GSPSVQCRS-----PRGNIGGKTL-----VSQ 174

232 VPPTFSRPETISEVMGLGNLNLCTAVSGPMPHVMMKGSBDLTENEMPIGRNVLQ 291

175 ELQDSGTWTCTVLQNGKVEFKIDIVPRASALPAPT 211

Db 292 NIGSANTCTAAGTGGID-SVSVV-KVQSLPRAFT 326

RESULT 40

S34338

biliary glycoprotein P - mouse

N/Alternate names: mouse hepatitis virus (MHV) receptor glycoprotein

C/Species: Mus musculus (house mouse)

C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999

C/Accession: S34338, J01510, A41093

R/Huang, D.C.; Huang, X.F.; Novel, M.; Novel, G.

Submitted to the EMBL Data Library, July 1992

A/Description: A Clp-family gene present on the lactose-protease plasmid of lactococcus

A/Reference number: S34338

A/Accession: S34338

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-521 <HVA>

A/Cross-references: EMBL:X67281; NID:G312585; PIDN:CAA47698.1; PID:G312586

R/McQuay, K.; Rosenberg, M.; Nedellec, P.; Turbide, C.; Beauchemin, N.

Gene 127, 173-183, 1993

A/Title: Expression of the Bgp gene and characterization of mouse colon biliary glycopro

A/Reference number: J01505; MUID:93273228; PMID:8500759

A/Accession: J01510

A/Molecule type: mRNA

A/Residues: 1-81, 'Q', 83-141, 'P', 143-521 <MCC>

A/Cross-references: GB:X67281

R/Williams, R.K.; Jiang, G.S.; Holmes, K.V.

Proc. Natl. Acad. Sci. U.S.A. 88, 5533-5536, 1991

A/Title: Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen

A/Reference number: A41093; MUID:91286498; PMID:1648219

A/Accession: A41093

A/Status: preliminary

A/Molecule type: protein

A/Residues: 35-59 <WIL>

C/Comment: This protein is expressed at the cell surface and plays a determinant role in

C/Genetics:

A/Gene: Bgpf

C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi

C/Keywords: glycoprotein; receptor

F/1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>

F/160-219/Domain: immunoglobulin homology <IMM1>

F/254-303/Domain: immunoglobulin homology <IMM2>

F/339-386/Domain: immunoglobulin homology <IMM3>

F/87,104,148,199,206,210,226,258,290,294,304,333,375/Binding site: carbohydrate (Asn) (C

Query Match 8.4%; Score 110; DB 2; Length 521;

Best Local Similarity 23.7%; Pred. No. 1.4;

Matches 45; Conservative 34; Mismatches 69; Indels 42; Gaps 10;

Query 13 VLGLALPAAATGKRVVGGKKGVTELTCTA-SQKSIQFHKNSNOIKILGNQGSFLTK 71

Db 140 VHQLLKKNITSNNSNPV-EGDDSVSLTCDSDYDPDNTYLSMRN-----GESLSE 189

Query 72 GPKSLNDRADRRRLMOCNFPILIKNLKIDSDTYICEV-----DQKEVQL-LVFG- 124

Db 190 GD-----RLKLSGNRTLTLLANTNDTPYCEETNPVSVKRSDFSNITTYGP 239

Query 125 ---LTANSPTHLGGQSLTLTLSPGSSPSVQC-----RSPRGNIGQKTLISVSOLE 176

Db 240 DTPILSPEDYILHFGSNLNLSCA--ASNPAQYFWLINEPHASS-----QELFIPMITT 293

Query 177 QDSGTWTCTV 186

Db 294 NNSGTYTCLV 303

RESULT 41

JN0581

vascular cell adhesion molecule-1 long splice form precursor - mouse

C/Species: Mus musculus (house mouse)

C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999

C/Accession: JN0581; J06074; A40275; A48919

R/Araki, M.; Araki, K.; Vassalli, P.

Gene 126, 261-264, 1993

A/Title: Cloning and sequencing of mouse VCAM-1 cDNA.

A/Reference number: JN0581; MUID:93246254; PMID:7683304

A/Accession: JN0581

A/Molecule type: mRNA

A/Residues: 1-739 <ARA>

A/Cross-references: EMBL:X67783; NID:G298116; PIDN:CAA47989.1; PID:G298117

R/Hession, C.; Moy, P.; Tizard, R.; Christou, P.; Williams, C.; Wyse, M.; Burdly, L.; M

Biochem. Biophys. Res. Commun. 183, 163-169, 1992

A/Title: Cloning of murine and rat vascular cell adhesion molecule-1.

A/Reference number: J06074; MUID:92181437; PMID:1371918

A/Accession: J06074

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-739 <HES>

A/Cross-references: GB:M84487; NID:G202345; PIDN:AAA40545.1; PID:G202346

R/Miyake, K.; Medina, K.; Ishihara, K.; Kimoto, M.; Auerbach, R.; Kinrade, P.W.

J. Cell Biol. 114, 557-565, 1991

A/Title: A VCAM-like adhesion molecule on murine bone marrow stromal cells mediates bin

A/Reference number: A40275; MUID:91317872; PMID:1713592

A/Accession: A40275

A/Molecule type: protein

A/Residues: 'XX', 27-32 <MIY>

R/Cybulsky, M.I.; Allan-McMead, M.; Collins, T.

Genomics 18, 387-391, 1993

A/Title: Structure of the murine VCAM1 gene.

A/Reference number: A48919; MUID:94117008; PMID:7507076

A/Accession: A48919

A/Molecule type: DNA

A/Residues: 1-692, 'N', 694-739 <CVB>

A/Cross-references: GB:I22355; NID:G47981; PIDN:AAA16921.1; PID:G459893; GB:I22301

C/Comment: This protein is a transmembrane protein and interacts with the beta-1 integr

C/Superfamily: immunoglobulin homology

C/Keywords: alternative splicing; cell adhesion; glycoprotein; transmembrane protein

F/1-24/Domain: signal sequence #status predicted <SIG>

F/25-739/Product: vascular cell adhesion protein-1 #status predicted <MAT>

F/239-293/Domain: immunoglobulin homology <IMM1>

F/338-385/Domain: immunoglobulin homology <IMM2>

F/527-581/Domain: immunoglobulin homology <IMM3>

F/659-720/Domain: transmembrane #status predicted <TMN>

F/225,273,424,531,561/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.4%; Score 110; DB 2; Length 739;

Best Local Similarity 21.3%; Pred. No. 2.2;

Matches 40; Conservative 34; Mismatches 68; Indels 46; Gaps 6;

Query 32 KKGPVTELTCTAQSQKSIQFHW--KNSNOI-KILGNQGSFLTKGPKSLNDRADRRSLMD 88

Db 237 QEGAVTMTCSSEGLPAPBLFGKRLDNEVQL----- 270

Query 89 QGNFPLIKNLKIEDSDTYICE-----VEDQKEVQLV-----FGLTANSPTHLGGQSL 138

Db 271 SGNAITLILMRNEDSVVYCEGNILGRKAEVELVQKPFIVDISPSQVAAQVGD 330

Query 139 LTLTLESPGSSPSVQCSPRGNIGQ-----GKTLISQLELDQSGTWTCTVLOXK 191

Db 331 VVLTCAIIGDSSPSFWRITDTPLVNGVNVNEGAKSTLVLSVGFEDSHSYLCAVTCLOR 390

Query 192 KVEFKIDI 199

Db 391 TLEKRTQV 398

RESULT 42

S04693

T-cell receptor delta chain (NTD4) - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-Jul-1999

C/Accession: S04693

R/Yoshikai, Y.; Matsuzaki, G.; Takeda, Y.; Ohga, S.; Kishihara, K.; Yuuki, H.; Nomoto,

Eur. J. Immunol. 18, 1039-1043, 1988

A>Title: Functional T cell receptor delta chain gene messages in athymic nude mice.
 A/Reference number: S04693; MUID:88296650; PMID:2969817
 A/Accession: S04693
 A/Molecule type: mRNA
 A/Residues: 1-279 <YOS>
 A/Cross-references: EMBL:X12729; NID:954894; PIDN:CAA1222.1; PID:954895
 A/Note: this sequence was determined from the differentially expressed gene
 C/Superfamily: Immunoglobulin V region; immunoglobulin homology
 C/Keywords: T-cell receptor

Query Match 8.3%; Score 109.5; DB 2; Length 279;
 Best Local Similarity 21.8%; Pred. No. 0.75;
 Matches 50; Conservative 29; Mismatches 85; Indels 65; Gaps 8;

QY 65 QGSEFLTGPSKLNDRADSRSLMDQGNPILIKNKIEDSYTC---EVEDQKEVOL 120
 DB 67 QGSEFLTGPSKLNDRADSRSLMDQGNPILIKNKIEDSYTC---EVEDQKEVOL 113
 121 LVFGLTANSDTHLLQGSLTLTLSPSSSPVQCRPRGKNI----- 163
 DB 114 LVFGLTANSDTHLLQGSLTLTLSPSSSPVQCRPRGKNI----- 162
 QY 164 -----OGKTLVSQLELDSDGTCTVTLONOKKVEKIDIVPRASALPA 208
 DB 163 SSKKIVFDPRAIVISPGKXSAVVLGGYGDNSVTCSVCHNSRTVH-STDFEYANSF-- 219
 QY 209 PPGSALPDPQTSALPDPAPASALPALAVISFLGLGACVACTART 257
 DB 220 --NNKKLPEPNDPTQISEPCGPRVIVHTEKVMMSLTVLGLRLFAKT 266

RESULT 43

S04663
 T-cell receptor gamma chain - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000
 C/Accession: S04663
 R/Okada, A.; Bank, I.; Rogozinski, L.; Takihara, Y.; Mak, T.W.; Chees, L.; Alt, F.W.
 J. Exp. Med. 168, 1481-1486, 1988
 A/Title: Structure of the gamma/delta T cell receptor of a human thymocyte clone.
 A/Reference number: S04663; MUID:89010543; PMID:2844954
 A/Accession: S04663
 A/Molecule type: DNA
 A/Residues: 1-304 <OKA>
 A/Cross-references: EMBL:X15018
 A/Note: the authors translated the codon AAC for residue 72 and AAT for residues 102 and
 C/Genetics:
 A/Intons: 15/1; 133/3
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: T-cell receptor
 F/158-223/Domain: immunoglobulin homology <IMM>

Query Match 8.3%; Score 109.5; DB 2; Length 304;
 Best Local Similarity 23.0%; Pred. No. 0.83;
 Matches 47; Conservative 29; Mismatches 61; Indels 67; Gaps 9;

QY 8 RHLLVVLQALLPRAATGKNKVLGKGGDT-----VLLTCTASAKSKISQCHM----- 53
 DB 2 RMALLLVLLAFSPASQSSNLEGGTKSVTRPRTSSAETCDLTVINAFYIHWYLAHQGKA 61
 QY 54 -----KNSNDIKILG-NOGSFLTKGPSKLNDRADSRSLMDQGNPILIKNKIED 103
 DB 62 PQRLLYDVNSKSKVLESGLSPGKRYTHTP-----RRMSV-----ILILRLNLEND 107
 QY 104 SDTYICEVDQKEVOLLVFGLTANSDTHLLQGSLTLTL--LESPPSSSPVQCRSPRG 160
 DB 108 SGVYCATWDRQDK--KLFG-----SGTTLVVTQKQADADAVSPKPTIFLP----- 150
 QY 161 KNIQGGKTLVSQLELDSDGTCTVT 184
 DB 151 -----STAFKTKQAKAGTYLC 165

RESULT 44

A41060
 neural cell adhesion molecule L1 precursor - human
 N/Alternate names: L1CAM
 C/Species: Homo sapiens (man)
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C/Accession: A41060; S18454; A35331; S21971; S21972; A60223; A31072; G02506
 R/Hlavin, M.L.; Lemmon, V.
 Genomics 11, 416-423, 1991
 A/Title: Molecular structure and functional testing of human L1CAM: an interspecies com
 A/Reference number: A41060; MUID:92120663; PMID:1769655
 A/Accession: A41060
 A/Molecule type: mRNA
 A/Residues: 1-1257 <HLA>
 A/Cross-references: GB:M64296; NID:9186053; PIDN:AAC14352.1; PID:93068548
 R/Kobayashi, M.; Miura, M.; Asou, H.; Uyemura, K.
 Biochim. Biophys. Acta 1090, 238-240, 1991
 A/Title: Molecular cloning of cell adhesion molecule L1 from human nervous tissue: a com
 A/Reference number: S18454; MUID:92031698; PMID:1932117
 A/Accession: S18454
 A/Molecule type: mRNA
 A/Residues: 1-3, 'V', 5-215, 'T', 217-249, 'T', 251-275, 'SV', 278-356, 'E', 358-625, 'V', 627-1257
 A/Cross-references: EMBL:X59847; NID:935009; PIDN:CAA42508.1; PID:935010
 A/Note: the authors translated the codon GAA for residue 27 as Gly
 R/Djabali, M.; Matrei, M.G.; Nguyen, C.; Roux, D.; Demengeot, J.; Denizot, F.; Moos, M.,
 Genomics 7, 587-593, 1990
 A/Title: The gene encoding L1, a neural adhesion molecule of the immunoglobulin family,
 A/Reference number: A35331; MUID:90353957; PMID:238785
 A/Accession: A35331
 A/Molecule type: DNA
 A/Residues: 332-371 <DJAB>
 A/Cross-references: GB:M55271
 R/Rosenthal, A.; Mackinnon, R.N.; Jones, D.S.C.
 Nucleic Acids Res. 19, 5395-5401, 1991
 A/Title: PCR walking from microdissection clone M54 identifies three exons from the hum
 A/Reference number: S21971; MUID:92020233; PMID:1923824
 A/Accession: S21971
 A/Molecule type: DNA
 A/Residues: 1082-1176 <ROS>
 A/Cross-references: EMBL:X58775; NID:929642; PIDN:CAA1576.1; PID:929643
 A/Accession: S21972
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: mRNA
 A/Residues: 353-935, 'V', 937-1176 <RO2>
 A/Cross-references: EMBL:X58776; NID:929644; PIDN:CAA37831.1; PID:94467833
 R/Hatper, J.R.; Prince, J.T.; Healy, P.A.; Stuart, J.K.; Nauman, S.J.; Scallcup, W.B.
 J. Neurochem. 56, 797-804, 1991
 A/Title: Isolation and sequence of partial cDNA clones of human L1: homology of human a
 A/Reference number: A60223; MUID:9132183; PMID:1993895
 A/Accession: A60223
 A/Status: not completed with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1030-1115, 'WLC', 1118-1176, 1181-1257 <HAR>
 R/Moff, J.M.; Frank, R.; Mujo, K.; Spiro, R.C.; Reisfeld, R.A.; Rathjen, F.G.
 J. Biol. Chem. 263, 11943-11947, 1988
 A/Title: A human brain glycoprotein related to the mouse cell adhesion molecule L1.
 A/Reference number: A31072; MUID:88298876; PMID:3136168
 A/Accession: A31072
 A/Molecule type: protein
 A/Residues: 'Q', 21-36 <WOL>
 R/Platzner, M.; Bauer, D.; Drescher, B.
 submitted to the EMBL Data Library, March 1995
 A/Reference number: H01368
 A/Accession: G02506
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1257 <PLA>
 A/Cross-references: EMBL:U52112; NID:91302657; PIDN:AAC51746.1; PID:91302658
 C/Genetics:
 A/Gene: GDB:L1CAM
 A/Cross-references: GDB:120133; OMIM:303350; OMIM:308840
 A/Map position: Xq28-Xq28
 A/Intons: 26/1; 31/1; 66/2; 134/1; 175/1; 232/1; 269/2; 331/1; 375/1; 460/2; 51

/2
C/Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology; i
C/Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; transmembran
P:1-19/Domain: signal sequence #status predicted <SIG>
P:20-125/Product: neural cell adhesion molecule L1 #status predicted <MAM>
P:257-314/Domain: immunoglobulin homology <IMM1>
P:533-593/Domain: immunoglobulin homology <IMM2>

Query March 8.3%; Score 109; DB 1; Length 1257;

Best Local Similarity 20.9%; Pred. No. 4.8;

Matches 48; Conservative 35; Mismatches 97; Indels 50; Gaps 9;

QY 8 RHLLLV-----LQALLPAA--ATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNQKJ 61

DB 409 RHGLLAAAYIVVQLPKILTDQNTYMAVQSGFAYLLCKAFGAPVSVQMLDQDTTV 468

QY 62 LGNQGSLITKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEV 121

DB 469 LQDERFF-----PYANGFLGIRDLQANDTGRYFCLANDQNNVTIM 509

QY 122 VFGLTANDTHLQ-----GSLITLES--PRSSSPVQCRSFRKNIQ----- 164

DB 510 A-NLKVKADATQITQGPSTTEKKSGRVFTQCAFDPGLQPSITWRGD-GRDLQELGSD 567

QY 165 -----GKTLVSQLELDQSGTWTCTVLQNGKVEFKIDIVPRASALPAP 209

DB 568 KYFIEDGR-LVTHSLDYSDQGNVSCVASTELDVESRAQLVGSFGFVP 616

RESULT 45

A41228
protein-tyrosine kinase (EC 2.7.1.112) Flk-1 precursor, endothelial cell-specific recept

C/Species: Mus musculus (house mouse)

C/Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 04-Feb-2000

C/Accession: A41228; A46065; I56365; S18832; S29991

C/Author: W. J. Jordan, C.T. Gavin, M. Jenkins, N.A. Copeland, N.G. Lemischka, I.R. Proc. Natl. Acad. Sci. U.S.A. 88, 9026-9030, 1991

A/Title: A receptor tyrosine kinase cDNA isolated from a population of enriched primitiv

A/Reference number: A41228; MUID:92020584; PMID:1717995

A/Accession: A41228

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1367 <MAM>

A/Cross-references: GB:X59397; NID:G50976; PIDN:CAA42040.1; PID:G50977

R:Miller, B.; Witzmann-Voos, S.; Schumacher, H.; Martinez, R.; Moller, N.P.; Risseu, W.; Cell 72, 835-846, 1993

A/Title: High affinity VEGF binding and developmental expression suggest Flk-1 as a maj

A/Reference number: A46065; MUID:93208880; PMID:7681362

A/Accession: A46065

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-74, 'T', '26-782, 'VL', '785-916, 'C', '918-1367 <MIL>

A/Cross-references: GB:X70842; NID:G57923; PIDN:CAA50192.1; PID:G57924

A/Note: submitted to the EMBL Data Library, January 1993

A/Note: sequence extracted from NCBI backbone (NCBI:128066)

R:Oelrichs, R.B.; Reid, H.H.; Bernard, O.; Ziemlecki, A.; Wilks, A.F. Oncogene 8, 11-18, 1993

A/Title: NTK/Flk-1: a putative receptor protein tyrosine kinase isolated from E10 embry

A/Reference number: 158365; MUID:93141255; PMID:8423988

A/Accession: 158365

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-578, 'D', '680-1340, 'RSPV', <OEL>

A/Cross-references: GB:S53103; NID:G264004; PIDN:AAB25043.1; PID:G264005

C/Genetics: C/Gene: FLK-1; NYK

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C/Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; transmembrane

P:830-1165/Domain: protein kinase homology <KIN>

P:83-846/Region: protein kinase ATP-binding motif

Query March 8.3%; Score 109; DB 2; Length 1367;

Best Local Similarity 21.0%; Pred. No. 5.3;

Matches 45; Conservative 36; Mismatches 73; Indels 60; Gaps 9;

QY 9 HLLVLQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNQKJ 68

DB 544 HVIGPEITVQPAQPTQ-----BSVSLCTADRTFENLTWK-----LGSQYTS 590

QY 69 LTKPSKLNDRADSRSLM-----DQGNFPLI--KNKIEDSDTYICEVEDQKEE 117

DB 591 VHMEES-LTPVCKRDLALMKNLGMFNSNTDILIVAFQNASLDDQGYVCSADKTKK 649

QY 118 -----VQLVFGLTANDTHLQGSILITLESPPGSSPVQCRSPRKN----- 162

DB 650 RHCLVKQILIERMAPITGNLENQTTI-----GETIEVTC--PASGNPTPIITWKD 701

QY 163 -----IQGKTLVSQLELDQSGTWTCT 164

DB 702 NETIVEDSGIVLRDGNRLTIRVRKEDGLYTC 735

RESULT 46

A31923
amalgam protein precursor - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 21-Jul-2000

C/Accession: A31923

R:Seeger, M.A.; Haefliger, L.; Kaufman, T.C. Cell 55, 589-600, 1988

A/Title: Characterization of amalgam: a member of the immunoglobulin superfamily from D

A/Reference number: A31923; MUID:89028670; PMID:3141062

A/Accession: A31923

A/Molecule type: DNA

A/Residues: 1-333 <SSE>

A/Cross-references: GB:M2361; NID:G156920; PIDN:AAA28367.1; PID:G156921

C/Genetics: C/Gene: FlyBase:Ana

A/Cross-references: FlyBase:FBgn0000071

Query March 8.2%; Score 108.5; DB 2; Length 333;

Best Local Similarity 20.8%; Pred. No. 1.1;

Matches 49; Conservative 38; Mismatches 90; Indels 59; Gaps 10;

QY 10 LLLVLQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHM-----KNSNQ 58

DB 10 LIFCLALSLSVLSAPYISQSKVAVSGDSVFNCIVEVQGLSVMARPSBSPDTS 69

QY 59 I-----KILG---NQGSLITKGPSKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYIC 109

DB 70 VLISMRNLLSLPDRVAVVTVEGPKT-----GSAIYFRIONIEVSDMGPEYC 117

QY 110 EV-----EQKEEVQLL--VFGLTANDTHLQGSILITLESPPGSSPV----- 153

DB 118 QVLVSATEKVKKLSIDIKTPPVIAENTPKSTLVTEQNLTECHANGPPKPTISWAREH 177

QY 154 QCRSPRGNITQGGTILSVQLELDQSGTWTCTVLQON-----OKVVEFKIDI 199

DB 178 NAWVPAGHLLAETTLIRSVHRMDRGYYC-IAQNEGQPDRLIRVEVFRFQI 232

RESULT 47

S00682
IGE FC receptor alpha chain precursor - human

N/Alternate names: Fc-epsilon receptor

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 23-Jul-1999

C/Accession: S00682; B30154; S42209

R:Kochan, J.; Pertine, L.F.; Hakimi, K.; Kline, J.P. Nucleic Acids Res. 16, 3584, 1988

A/Title: Isolation of the gene coding for the alpha subunit of the human high affinity

A/Reference number: S00682; MUID:88233953; PMID:2967464

A/Accession: S00682

A/Molecule type: mRNA

A/Residues: 1-257 <KOC>

A/Cross-references: EMBL:X06948; NID:G31317; PIDN:CAA30025.1; PID:G31318

R.Shimizu A.; Tepler, I.; Benfey, P.N.; Berenslein, E.H.; Siraganian, R.P.; Leder, P.
Proc. Natl. Acad. Sci. U.S.A. 85, 1907-1911, 1988
A>Title: Human and rat mast cell high-affinity immunoglobulin E receptors: characterized
A/Reference number: A94191; MUID:88158102; PMID:2964640
A/Accession: B30154
A/Molecule type: mRNA
A/Residues: 1-257 <SH1>
A/Cross-references: GB:J03605; NID:g187449; PIDN:AAA36204.1; PID:g307164
R.Yagi, S.; Yanagida, M.; Tanida, I.; Hasegawa, A.; Okumura, K.; Ka, C.
Eur. J. Biochem. 220, 593-598, 1994
A>Title: High-level expression of the truncated alpha chain of human high-affinity receptor product.
A/Reference number: S42209; MUID:94170811; PMID:8125119
A/Accession: S42209
A/Molecule type: protein
A/Residues: 26-197 <YAG>
A/Experimental source: purified recombinant protein
C/Genetics:
A/Gene: GDB:FCER1A
A/Cross-references: GDB:119902; OMIM:147140
A/Map position: 1q23-1q23
C/Superfamily: FC gamma receptor III; immunoglobulin homology
C/Keywords: immunoglobulin receptor; transmembrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-257/Product: IGE FC receptor alpha chain #status predicted <MAT>
F:44-95/Domain: immunoglobulin homology <IMM1>
F:125-178/Domain: immunoglobulin homology <IMM2>

Query Match 8.2%; Score 108; DB 2; Length 257;
Best Local Similarity 27.1%; Pred. No. 0.88;
Matches 55; Conservative 22; Mismatches 64; Indels 62; Gaps 10;

OY 10 LLLVQLALP-----AATGKRVLGK-----KGTVELTCTASQKSIQFHWKNSNQI 59
DB 10 LLLVQLALP-----AATGKRVLGK-----KGTVELTCTASQKSIQFHWKNSNQI 51
OY 60 KIIGNGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVE--DQKEE 117
DB 52 ----NGNPFVSSSTKPFHNG----SLSEETNSSLNINAKFEDSGYKCKQHQQVNESEP 103
OY 118 VOLLVFGLTANSDTHLQGSLLTLTLESPGSSPSVOCRSPPRGKNI-----QGKTL-- 169
DB 104 VYLEVF-----SDMLLLQA--SAEYVMEGQF---LFLKCHGMWMDVYKVIYVDGELTKY 154
OY 170 -----SVSOLELQDSGTWTCT 185
DB 155 WYENHNISITNATVEDSGTYTCT 177

RESULT 48
S40370
IG kappa chain - human
C/Species: Homo sapiens (man)
C/Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C/Accession: S40370
R.Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A>Title: Expressed human immunoglobulin chi genes and their hypermutation.
A/Reference number: S40312; MUID:94080891; PMID:8258341
A/Accession: S40370
A/Status: preliminary; translation not shown
A/Molecule type: mRNA
A/Residues: 1-122 <KLE>
A/Cross-references: EMBL:X72480; NID:g441428; PIDN:CAA51148.1; PID:g441429
C/Superfamily: immunoglobulin V region; immunoglobulin homology
C/Keywords: heterocramer; immunoglobulin
F:30-104/Domain: immunoglobulin homology <IMM>

Query Match 8.2%; Score 107.5; DB 2; Length 122;
Best Local Similarity 35.5%; Pred. No. 0.39;
Matches 38; Conservative 5; Mismatches 53; Indels 11; Gaps 3;

OY 11 LVLVQLALPAA-----TQGNKVVLGKGDVLELTCTASQKSIQFHWKNSNQIKILGN 64

DB 1 LGLLLMLLRGASCDIQLTQSSSLASVGDVTTTCRASQSISTFLHWQON----LGK 56
OY 65 QGSFLLTKGPKSLNDRADSRSLMDQGNFPLIKNIKIEDSDTYICE 110
DB 57 APRLLIYASNLQSGVPSRFSGSGSDTFTLTISGLQPDFAFYCYQ 103

RESULT 49
IHHUNG
neural cell adhesion molecule 1 GPI-anchored splice form precursor, muscle-specific - h
N/Alternate names: CD56; NCAM-120
C/Species: Homo sapiens (man)
C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 28-Jan-2000
C/Accession: S07784; A26883
R/Barton, C.H.; Dickson, G.; Gower, H.J.; Rowett, L.H.; Putt, W.; Elsom, V.; Moore, S.E
Development 104, 165-173, 1988
A>Title: Complete sequence and in vitro expression of a tissue-specific phosphatidylinositol
A/Reference number: S07784; MUID:89305258; PMID:3253057
A/Accession: S07784
A/Molecule type: mRNA
A/Residues: 1-761 <BAR>
A/Cross-references: EMBL:X16841; NID:g35005; PIDN:CAA34739.1; PID:g35006
R/Dickson, G.; Gower, H.J.; Barton, C.H.; Prentice, H.M.; Elsom, V.L.; Moore, S.E.; Cox, Cell 50, 1119-1130, 1987
A>Title: Human muscle neural cell adhesion molecule (N-CAM): identification of a muscle
A/Reference number: A90895; MUID:87301755; PMID:2887295
A/Accession: A26883
A/Molecule type: mRNA
A/Residues: 491-761 <DIC>
A/Cross-references: GB:M17409; NID:g189097; PIDN:AAA59912.1; PID:g386979
C/Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM moi
C/Genetics:
A/Gene: GDB:NCAM1; NCAM; CD56
A/Cross-references: GDB:119448; OMIM:116930
A/Map position: 11q22.2-11q22.3
C/Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; imm
C/Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane
F:1-15/Domain: signal sequence #status predicted <SIG>
F:20-761/Product: neural cell adhesion molecule phosphatidylinositol-linked form, muscl
F:34-98/Domain: immunoglobulin homology <IMM1>
F:132-191/Domain: immunoglobulin homology <IMM2>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-289/Domain: immunoglobulin homology <IMM3>
F:263-272/Region: NCAM binding #status predicted
F:322-387/Domain: immunoglobulin homology <IMM4>
F:419-481/Domain: immunoglobulin homology <IMM5>
F:499-587/Domain: fibronectin type III repeat homology #status atypical <FN3A>
F:633-720/Domain: fibronectin type III repeat homology #status atypical <FN3B>
F:41-96, 139-189, 235-287, 329-385, 426-479/Disulfide bonds: #status predicted
F:222, 315, 347, 423, 449, 478/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.1%; Score 107; DB 1; Length 761;
Best Local Similarity 19.3%; Pred. No. 3.7;
Matches 44; Conservative 38; Mismatches 102; Indels 44; Gaps 7;

OY 3 RG-VPRHLLVQLALPAAATGKRVLGKGDVLELTCTASQKSIQFHWKNSNQIKI 61
DB 196 RGEINFKDIOIVNVPPITQARQINATANIGQSVTLVCDAGPEPTMSW----- 247
OY 62 LGNGSFLTKGPKSLNDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVE--DQKEE 117
DB 248 -----TKDGEIQIEBDEKTYIFSDSSQLTIKVKKNDAEYICIAENKAGQDAT 299
OY 118 VOLLVFGLTANSDTHLQGSLLTLTLESPGSSPSVOCRSPPRGKNIQGG-- 166
DB 300 IHLKVFAPKPTIVENQTMLEBQVTLTCEASGDPFISTWTSTRNISSEKTLDDGM 359
OY 167 -----KTLVSQLELQDSGTWTCT---TVLQNGKRVKFEKIDIVPR 202
DB 360 VVSHARVSSLTLSKIQTVDAGYICTASNTIGDSQSMYLEVQVAPK 407

```

RESULT 50
S18252
heparan sulfate proteoglycan - mouse
N/Alternate names: perlecan
C/Species: Mus musculus (house mouse)
C/Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 05-Nov-1999
C/Accession: S18252; A31917; B31917; S66460
R/Nonan, D.M.; Fuller, A.; Valente, P.; Cai, S.; Horigan, E.; Sasaki, M.; Yamada, Y.; Ha
J. Biol. Chem. 266, 22939-22947, 1991
A/Title: The complete sequence of perlecan, a basement membrane heparan sulfate proteogly
adhesion molecule.
A/Reference number: S18252; MUID:92078153; PMID:1744087
A/Accession: S18252
A/Molecule type: mRNA
A/Residues: 1-3707 <NO>
A/Cross-references: EMBL:M77174; NID:G200295; PIDN:AAA39911.1; PID:G200296
R/Nonan, D.M.; Horigan, E.A.; Ledbetter, S.R.; Vogeli, G.; Sasaki, M.; Yamada, Y.; Hase
J. Biol. Chem. 263, 16379-16387, 1988
A/Title: Identification of cDNA clones encoding different domains of the basement membra
A/Reference number: A32680; MUID:89034110; PMID:2972708
A/Accession: A31917
A/Molecule type: mRNA
A/Residues: 940-1601 <NO2>
A/Cross-references: GB:J04054; NID:G200252; PIDN:AAA39899.1; PID:G200253
A/Accession: B31917
A/Molecule type: mRNA
A/Residues: 1870-2600 <NO3>
A/Cross-references: GB:J04055; NID:G200300; PIDN:AAA39912.1; PID:G200301
R/Schulze, B.; Mann, K.; Bartschke, R.; Wiedemann, H.; Timpl, R.
Eur. J. Biochem. 231, 551-556, 1995
A/Title: Structural properties of recombinant domain III-3 of perlecan containing a glob
A/Reference number: S66460; MUID:95377282; PMID:7649154
A/Accession: S66460
A/Molecule type: protein
A/Residues: 1272-1274, 'X', 1276, 'X', 1278-1279 <SCH>
C/Superfamily: LDL receptor ligand-binding repeat homology; EGF homology; laminin G repe
C/Keywords: glycoprotein
F/199-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F/283-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F/325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F/368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F/764-811/Domain: laminin-type EGF-like homology <LEG>
F/1159-1206/Domain: laminin-type EGF-like homology <LEG7>
F/1563-1610/Domain: laminin-type EGF-like homology <LEG8>
F/1613-1668/Domain: laminin-type EGF-like homology <LEG8>
F/3163-3198/Domain: EGF homology <EGF>
F/3270-3423/Domain: laminin G repeat homology <LG2>
F/3464-3492/Domain: EGF homology <EGF7>
F/1256,1891,2336,2394,2427/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 8.1%; Score 107; DB 2; Length 3707;
Best Local Similarity 21.7%; Pred. No. 24;
Matches 55; Conservative 39; Mismatches 95; Indels 64; Gaps 10;

```

```

DB 2661 PEKAVRNQRLRL 2673
RESULT 51
A38096
perlecan precursor - human
N/Alternate names: basement membrane heparan sulfate proteoglycan; heparan sulfate prot
C/Species: Homo sapiens (man)
C/Date: 07-Apr-1994 #sequence revision 07-Apr-1994 #text_change 05-Nov-1999
C/Accession: A38096; S19256; S77946; A41059; A40306; B33625; A33625; A41736
R/Murdoch, A.D.; Dodge, G.R.; Cohen, I.; Tuan, R.S.; Iozzo, R.V.
J. Biol. Chem. 267, 8544-8557, 1992
A/Title: Primary structure of the human heparan sulfate proteoglycan from basement memb
tor, laminin, neural cell adhesion molecules, and epidermal growth factor.
A/Reference number: A38096; MUID:92235084; PMID:1569102
A/Accession: A38096
A/Molecule type: mRNA
A/Residues: 1-4391 <MUR>
A/Cross-references: GB:M85289; NID:G184426; PIDN:AAA52700.1; PID:G184427
R/Kallunki, P.; Tyygvaasen, K.
J. Cell Biol. 116, 559-571, 1992
A/Title: Human basement membrane heparan sulfate proteoglycan core protein: a 467-kD pr
all adhesion molecules, and epidermal growth factor.
A/Reference number: A41736; MUID:92112994; PMID:1730768
A/Accession: S19256
A/Molecule type: mRNA
A/Residues: 1-57, 'D', 59-434, 'A', 436, 'FL', 438-449, 'Q', 451-502, 'A', 503-792, 'K', 794-908, 'R'
71-2979, 'H', 2981-2994, 'G', 2996-3167, 'T', 3169-3240, 'R', 3242-3426, 'R', 3428-3631, 'Q', 3633-
A/Cross-references: EMBL:X62515; NID:G29469; PIDN:CAA44373.1; PID:G29470
R/Kallunki, P.; Eddy, R.L.; Byers, M.G.; Kesäniemi, M.; Shows, T.B.; Tyygvaasen, K.
Genomics 11, 389-396, 1991
A/Title: Cloning of human heparan sulfate proteoglycan core protein, assignment of the
A/Reference number: A41059; MUID:92120660; PMID:1685141
A/Accession: A41059
A/Molecule type: mRNA
A/Residues: 'R', 892-908, 'R', 910-1101, 'L', 1103-1132, 'L', 1134-1221, 'L', 1223-1397 <KA2>
A/Cross-references: GB:S76436; NID:G243370; PIDN:AA21121.1; PID:G243371
R/Dodge, G.R.; Kovalesky, I.; Chu, M.L.; Haesell, J.R.; McBride, O.W.; Yi, H.F.; Iozzo,
Genomics 10, 673-680, 1991
A/Title: Heparan sulfate proteoglycan of human colon: partial molecular cloning, cellu
A/Reference number: A40306; MUID:91365376; PMID:1679749
A/Accession: A40306
A/Molecule type: mRNA
A/Residues: 1018-1405, 'G', 1407-1409, 'G', 1411-1465 <DOD>
A/Cross-references: GB:M64253; NID:G184424; PIDN:AAA52699.1; PID:G184425
R/Heremans, A.; van der Schueren, B.; De Cock, B.; Paulsson, M.; Cassiman, J.J.; van de
J. Cell Biol. 109, 3199-3211, 1989
A/Title: Matrix-associated heparan sulfate proteoglycan: core protein-specific monoclona
anes.
A/Reference number: A33625; MUID:90078352; PMID:2687294
A/Accession: B33625
A/Molecule type: protein
A/Residues: 1379-1384, 'X', 1386-1388, 'X', 1390-1398 <HE2>
A/Accession: A33625
A/Molecule type: protein
A/Residues: 2166-2171, 'X', 2173-2175, 'X', 2177-2185 <HE3>
A/Note: peptide potentially matches four different regions of sequence shown
C/Genetics:
A/Gene: GDB:HSRG2
A/Cross-references: GDB:126372; OMIM:142461
A/Map position: 1936.1-1936.1
C/Superfamily: LDL receptor ligand-binding repeat homology; EGF homology; laminin G rep
C/Keywords: chondroitin sulfate proteoglycan; glycoprotein; heparan sulfate; transmembr
F/1-21/Domain: signal sequence #status predicted <SIG>

```


F:22-4391/Product: perlecan #status predicted <MAT>
 F:22-193/Domain: I <DOM1>
 F:194-530/Domain: II <DOM2>
 F:199-534/Domain: LDI receptor ligand-binding repeat homology <LDL1>
 F:285-319/Domain: LDI receptor ligand-binding repeat homology <LDL2>
 F:325-359/Domain: LDI receptor ligand-binding repeat homology <LDL3>
 F:368-403/Domain: LDI receptor ligand-binding repeat homology <LDL4>
 F:531-1676/Domain: III <DOM3>
 F:1159-1206/Domain: laminin-type EGF-like homology <LEG>
 F:163-1610/Domain: laminin-type EGF-like homology <EG7>
 F:1613-1668/Domain: laminin-type EGF-like homology <LEG8>
 F:1677-3686/Domain: IV <DOM4>
 F:2007-2034/Domain: transmembrane #status predicted <TRM>
 F:3687-4391/Domain: V <DOM5>
 F:3845-3880/Domain: EGF homology <EGF1>
 F:3888-3921/Domain: EGF homology <EGF>
 F:3953-4106/Domain: laminin G repeat homology <LG2>
 F:4147-4175/Domain: EGF homology <EGF2>
 F:4149-4151/Region: motor neuron attachment (L-R-E) motif
 F:4299-4301/Region: motor neuron attachment (L-R-E) motif
 F:6571-76/Binding site: heparan sulfate (Ser) (covalent) #status predicted
 F:89,554,1755,2121,3072,3105,3279,3780,3836,4068/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:2995,3933,4179/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 8.1%; Score 107; DB 2; Length 4391;

Best Local Similarity 19.8%; Pred. No. 30; Mismatches 94; Indels 146; Gaps 11;

Matches 67; Conservative 31; Mismatches 94; Indels 146; Gaps 11;

QY 12 LVLQLALLPATQGNKVVLLGKGDVVELTCTASQKSIQPHWKNNOIKILGNQSFLLTK 71
 DB 3291 ILHVESPFATVPEHNASVQAGETVQLCLAGHTPLTFLQW-----SR 3334
 QY 72 GPKSLNDRASRSLSMDQGNPPLIKLKIEDSDTYICEVDO---KEEVQLVFG--- 124
 DB 3335 VGSSLPERATARNEL-----LHFERAPEDSGRCRYTNKVGSAEAPQLVQGP 3387
 QY 125 -LTRANS-----DTHLQGGSLTLE-----SPGSSPS 152
 DB 3388 SLRATSLPAGSTPVTQVTPQLTKISIGASVDFHCANVSDGDTQLRWFEGGQLPFGHSVQ 3447
 QY 153 -----VQCRSPRGKN----- 162
 DB 3448 DGVLRIGNLDPSCGTYICQAHGPMGKAKQASQVLQALPSVLINIRTSVQTVVGHAVE 3507
 QY 163 -----IQGKTLVSQLELQDSGTWTCTVQLQ---NQK 192
 DB 3508 FECLALGDPKROYTWKVGHLRPGIVQSGGVARIHAHELADAGQYRCTNNAAGTQSH 3567
 QY 193 VEFKIDIVPRASALPAP---PTGSALDPDPTASALPP 227
 DB 3568 VLLIVQALPQIS-MPQEVPRPAGSAAVFPCLASGTPTP 3604

RESULT 52

B37266

Ig kappa chain V region (2G8) - mouse

C/Species: Mus musculus (house mouse)

C/Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 21-Jan-2000

C/Accession: B37266

R/Ruff-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.

J. Biol. Chem. 266, 6607-6613, 1991

A/Title: Heavy and light chain variable region sequences and antibody properties of anti

A/Reference number: A38740; MUID:91177923; PMID:11706720

A/Accession: B37266

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-111 <RUF>

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F:19-93/Domain: immunoglobulin homology <IRM>

Query Match 8.1%; Score 106.5; DB 2; Length 111;
 Best Local Similarity 35.2%; Pred. No. 0.42;

Matches 32; Conservative 10; Mismatches 40; Indels 9; Gaps 3;

QY 23 TQGNKVVLLGKGDVVELTCTASQKSIQPHW---KNSNOIKILGNQSFLLTKPSKLNDR 79
 DB 8 TQSTKIMTSVSDRYSITCKASQDASNYLWYQOKPDGTVKILYASYSRTG---VPDR 64
 QY 80 ADSRSLMDQGNPPLIKLKIEDSDTYICE 110
 DB 65 FTGRGSGTD---FSFTISSVKAEDLAIVYCE 92

RESULT 53

I38740

Ig kappa chain V region (Py42) - mouse

C/Species: Mus musculus (house mouse)

C/Date: 18-Oct-1991 #sequence_revision 18-Oct-1991 #text_change 21-Jan-2000

C/Accession: I38740

R/Ruff-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.

J. Biol. Chem. 266, 6607-6613, 1991

A/Title: Heavy and light chain variable region sequences and antibody properties of anti

A/Reference number: A38740; MUID:91177923; PMID:11706720

A/Accession: I38740

A/Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr.

A/Molecule type: mRNA

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F:19-93/Domain: immunoglobulin homology <IRM>

Query Match 8.1%; Score 106.5; DB 2; Length 111;
 Best Local Similarity 35.2%; Pred. No. 0.42;

Matches 32; Conservative 11; Mismatches 39; Indels 9; Gaps 3;

QY 23 TQGNKVVLLGKGDVVELTCTASQKSIQPHW---KNSNOIKILGNQSFLLTKPSKLNDR 79
 DB 8 TQSTKIMTSVSDRYSITCKASQDASNYLWYQOKPDGTVKILYASYSRTG---VPDR 64
 QY 80 ADSRSLMDQGNPPLIKLKIEDSDTYICE 110
 DB 65 FTGRGSGTD---FSFTISSVKAEDLAIVYCE 92

RESULT 54

JC1239

Opioid-binding protein (clones S68 and SG13) - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 19-May-2000

C/Accession: JC1239

R/Lippman, D.A.; Lee, N.M.; Loh, H.H.

Gene 117, 249-254, 1992

A/Title: Opioid-binding cell adhesion molecule (OBAM)-related clones from a rat brain

A/Reference number: JC1238; MUID:92347701; PMID:1339369

A/Accession: JC1239

A/Molecule type: mRNA

A/Residues: 1-345 <LIP>

C/Cross-references: GB:M88710; NID:G203247; PIDN:AAA0859.1; PID:G203248; GB:M88711; NID:

C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi

C/Keywords: transmembrane protein

Query Match 8.1%; Score 106.5; DB 2; Length 345;
 Best Local Similarity 22.1%; Pred. No. 1.6;
 Matches 58; Conservative 43; Mismatches 101; Indels 61; Gaps 10;

QY 11 LVLQLALLPATQGNKVVLLG-----KKGDTVELTCTASQKSIQPHWKNNS 57
 DB 14 LVVVSRLFLFVPTGVPRSGDAPTPKANDVTVRQGSATLRCITIDRVLT-RVAMLNRS 72
 QY 58 QIKILGN-----QGSFLTKGSKLNDRASRSLSMDQGNPPLIKLKIEDSDTYICE 110
 DB 73 TIIYAGNDKMSIDPRVITLIVNTPTQ-----YSIMIQNDVDVDEGPTCS 116
 QY 111 VE-----DQKEVQLLVFG-----LTANSPTHLQGGSLTTLSPSPGSSPSVQCRSPRGKN 162

Db 117 VOTDNHPTKSRVHLIVQVPQIMNISDITVNEISSVTLCLALGPREPTVTHLSYKE 176
 Qy 163 IQG-----GKTLVSQLEQDSGTWTCVTLQ-----NCKVEFKIDIVPRASALPAPPGS 213
 Db 177 GCGFVSEBEVLEISDIKDSGSEYCSALNDVAPADVKKIVTVPYIS--KAKNTGV 234
 Qy 214 ALPDPQTASALPDPASALPAA 236
 Db 235 SVGOKGILSC-----EASAVPMA 252

RESULT 55

T34416

hypochemical protein P12F3.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34416

R:Fulton, B.; Wohldmann, P.

A:Description: The sequence of C. elegans cosmid P12F3.

A:Reference number: 221521

A:Accession: T34416

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2783 <FULL>

A:Cross-references: EMBL:U80022; PIDN:AC25886.1; GSPDB:GN00023; CESP:P12F3.2

C:Genetic8:

A:Map position: 5

A:Gene: CESP:P12F3.2

A:Introns: 45/3; 90/3; 451/3; 509/1; 2313/3; 2341/3; 2378/3; 2414/2; 2453/3; 2474/2; 252

Query Match 8.1%; Score 106.5; DB 2; Length 2783;

Best Local Similarity 22.6%; Pred. No. 19;

Matches 56; Conservative 31; Mismatches 84; Indels 77; Gaps 10;

Qy 28 VLLGKKDPTVELTCTASQKSIQFHWKNSNQIKILNGSFLTKGPSKLNDRADSRSLW 87

Db 1187 VVLKTAGTATFTQSYANPAQVVM-----LHNGKALQOTKSNYKRLF 1231

Qy 88 DQGNPPLIINKLKIEDSDTYICEVEDQEEV---QLLVFGLTANS----- 129

Db 1232 DDMATVLIENVTDLCCTYAVANNQCGDVHTSQTLTSSGSAKIAALPYFIIEKLP 1291

Qy 130 DTHLQGSLLTLESPPGSSP-----SVQCRPR-----GKNIQSGTLSSSQ 173

Db 1292 KINVVEG--ATLSIQADLNGSPRPEVWLKONSELVESDRIQMKCDGNVY---LTVRD 1345

Qy 174 LELDQSGTWCTCT-----VLQN-----QKVEFKIDIVPRASALPAPPGSA 214

Db 1346 VGLEDEGTYITTAENEKIKINQNTVEVSTYSKVEKKEKKVEKDGKKKPGRP--G 1402

Qy 215 LPDPQTAS 222

Db 1403 LPRPSGAS 1410

RESULT 56

T43027

neural cell adhesion molecule L1 - goldfish

N:Alternate names: E587 antigen

C:Species: Carassius auratus (goldfish)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000

C:Accession: T43027

R:Giordano, S.; Leesing, U.; Lottspeich, F.; Stuenkel, C.A.O.

A:Description: Molecular cloning of goldfish E587 antigen, a cell adhesion molecule expr

A:Reference number: 222294

A:Accession: T43027

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1232 <GIO>

A:Cross-references: EMBL:U55211; NID:G1305526; PID:G1305527; PIDN:AAA99159.1

C:Superfamily: neural cell adhesion molecule L1, fibronectin type III repeat homology;
 C:Keywords: cell adhesion; membrane protein

Query Match 8.0%; Score 106; DB 2; Length 1232;

Best Local Similarity 23.2%; Pred. No. 7.8;

Matches 54; Conservative 40; Mismatches 87; Indels 52; Gaps 11;

Qy 2 NRGVPRFHLVLVQLALLPATQGNKVLGKGGDVELTCTASQKSIQFHWKNSNQIKI 61

Db 286 SQGSVKHIVAVTVAA--PWTRRPNHLIARGETVALLDQAEIPIPNITW--SMNAPI 342

Qy 62 LGNGSFLTKGPSKLNDRADSRSLMDQGNPPLIINKLKIEDSDTYICEVEDQEEVQLT 121

Db 343 AGTD-----PPRRNHV---SSQTLTLTVQI--SDTAVYHVEATKHNIL 383

Qy 122 VFGLTNSDTHL--LQGSLLT--LTLESPPGSSPSVQCR--SPRK----- 161

Db 384 I-----NTHVHVVELPQIILTEDDLKYATEGGQTVLLQCRFTGSPQPRVWQITNSGPAL 438

Qy 162 -NIQGKT---LSVQLEQDSGTWTCVTLQNKQVEFKIDIVPRASALPAP 209

Db 439 ANAKMSQTSQNTLQISDVSEBSSMYTCSVSTMSISAEVLVNLTKIYDPP 491

RESULT 57

T08851

Down syndrome cell adhesion protein 1 - human (fragment)

N:Alternate names: Down syndrome cell adhesion molecule

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999

C:Accession: T08851

R:Yamakawa, K.; Huo, Y.K.; Haendel, M.A.; Hubert, R.; Chen, X.N.; Lyons, G.E.; Korenber

submitted to the EMBL Data Library, September 1997

A:Description: DSCAM: A novel member of the immunoglobulin superfamily maps in a down s

A:Reference number: 216495

A:Accession: T08851

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1896 <YAM>

A:Cross-references: EMBL:AF023449; NID:G3169765; PID:G3169766

A:Experimental source: brain; developmental stage: 14 weeks; fetal

C:Genetics:

A:Gene: DSCAM

A:Map position: 21q22

A:Note: derived from alternatively-spliced mRNA

C:Function: involved in nervous system development

A:Keywords: alternative splicing

Query Match 8.0%; Score 106; DB 2; Length 1896;

Best Local Similarity 21.5%; Pred. No. 13;

Matches 50; Conservative 34; Mismatches 75; Indels 74; Gaps 10;

Qy 34 GDTVELTCTASQKSIQFHWKNSNQIKILNGSFLTKGPSKLNDRADSRSLMDQGNP 93

Db 225 GQVLEPCXALGHPDPYRLKDNM-----FLELSGR-----FQKVTG 263

Qy 94 LIINKLKIEDSDTYICEVEDQEEVQLLVFGLTNSDTHLQGSLLTLT-----ESPFG 148

Db 264 LLIENIRPSDGSVCEVSNR-----YG-TAKVIGRLVYQPLKATISPRKVKSSVG 314

Qy 149 SSPVQCR-----SPRKNIQ-----GKTLVSQLEQDSGTWTCV 186

Db 315 SQVLSGCVTGTEDELQSWRNGEILNP--QKNVIRITGINNENLIMDMVNSDGGAYOCFV 373

Qy 187 LQNKQVEFKIDIVPRASALPAPPGSALPD--PQTASALPDPASALPAAL 237

Db 374 RKDLTSQDYQVYV-----LBDGPKIITISAEKVKVSPAEFVSL 412

RESULT 58

I49294

CD7 antigen - mouse

C;Species: Mus musculus (house mouse)
 C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
 C;Accession: I49294; I49588
 R;Lee, D.M.; Watson, M.L.; Seldin, M.F.
 Immunogenetics 39, 289-290, 1994
 A;Title: Mouse Cd7 maps to chromosome 11.
 A;Reference number: I49294; MUID:94164701; PMID:7509775
 A;Accession: I49294
 A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EHE
 A;Molecule type: DNA
 A;Residues: 1-210 <RES>
 A;Cross-references: EMBL:U23462; NID:9833815; PIDN:AB17482.1; PID:9833816
 R;Yoshikawa, K.; Sero, M.; Ueda, R.; Obata, Y.; Fukatsu, H.; Segawa, A.; Takahashi, T.
 Immunogenetics 37, 114-119, 1993
 A;Title: Isolation and characterization of mouse CD7 cDNA.
 A;Reference number: I49588; MUID:93138718; PMID:7678579
 A;Accession: I49588
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-68, 'F', 70-210 <RE2>
 A;Cross-references: GB:D10329; NID:93868752; PIDN:BA01171.1; PID:9286071
 C;Genetics:
 A;Gene: Cd7
 A;Intons: 25/1; 130/1; 175/3
 C;Keywords: glycoprotein

Query Match 8.0%; Score 105.5; DB 2; Length 210;
 Best Local Similarity 21.7%; Pred. No. 1;
 Matches 57; Conservative 33; Mismatches 60; Indels 113; Gaps 13;

10 LLLVLQIA-LLPA-----ATQGNKVLGKKDPVELTCTASQKSIQFMKNSNQIKL 62
 Db 7 LALLTLAAGLPLCPADQVHQSPRLTASBDSVNITCSR----- 48
 QY 63 GNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIKLIKIEDSPDYICEVEDQKEVOLV 122
 Db 49 GHLEGLIMK-----KIMPCA-----YNYVLYEDRQE----- 74
 QY 123 FGLTANSDTHLQGSLLTLESPPGSSPSVQCRSPRGK-NIQGK---TLSVSOLEPD 178
 Db 75 -----PTVD-RTFSGRINFSGSQKULTYITISLQAD 105
 QY 179 SGTWCTVLONOKVE---FKIDIVRASALPAPPGSALPDQTSALPDPPAAS-AL 233
 Db 106 TGDYTCAV---KRVASRGGLFTTVVNEKSS-----QEAIRSOEPLQTSFSF 149
 QY 234 PALAVISFLIGLIGLVACVLAR 256
 Db 150 PAALAVGFFFTGILLGVCSMLR 172

RESULT 59

connectin 3B - chicken (fragment)
 N;Alternate names: Cn3B protein
 C;Species: Gallus gallus (chicken)
 C;Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 17-Mar-1999
 C;Accession: P00568
 R;Maruyama, K.; Endo, T.; Kume, H.; Kawamura, Y.; Kanazawa, N.; Nakauchi, Y.; Kimura, S.;
 Biochem. Biophys. Res. Commun. 194, 1288-1291, 1993
 A;Title: A novel domain sequence of connectin localized at the 1 band of skeletal muscle
 A;Reference number: P00568; MUID:93356802; PMID:8352787
 A;Accession: P00568
 A;Molecule type: mRNA
 A;Residues: 1-1323 <MAR>
 A;Cross-references: DDBJ:D16541; NID:9391629; PID:d1004495; PID:9391630
 A;Experimental source: skeletal muscle
 C;Comment: This protein string-like single molecule spans from the 2 line to the M line

Query Match 8.0%; Score 105.5; DB 2; Length 123;
 Best Local Similarity 19.5%; Pred. No. 9.3;
 Matches 57; Conservative 45; Mismatches 98; Indels 93; Gaps 12;

QY 8 RHLLVLQIALLPATQGN-KVVLG-----KKGDTVELTCT 42
 Db 928 RHLLKIKNCQL---EDQGYRIYCGPHIASALTYIEPVERHLDPTTKRGNTCTLSQ 984
 QY 43 ASQKSIQPHMKNSNQIKLIGNQGSFLTGPSEKLNDRADSRSLMDQGNFPLIKLIKIE 102
 Db 985 FSLPNAKSGWYNNGRIKI---GGRYSTQVSKVH-----KLIKQVRE 1026
 QY 103 DSDTYICEVED-----QKEBVLVLFGLTANSDTHLQGSLLTLESPPG----- 148
 Db 1027 DQGYLYCKLDNLETADLTIEAEPIQF-----TKSIQNIWVSEHQSAFEECEVSPDDAV 1080
 QY 149 ----SSPSVQCRSPRGKNIQGGKT---LSVSOLEQDSGTWCT-----Y 186
 Db 1081 VTWYKGPTELTSPKSPFSEBGCWYTHNVTAEDEGVYSYIARLEPGEARSTAEVY 1140
 QY 187 LQNKQKVEFKIDIVPRAS-ALPAPPGSALPDQTSALPD---PAASALP 234
 Db 1141 VTKRILELKPDPDVPDAKVAVPPQKPAEAPIPILLPLPTPEKKKPAKKVP 1193

RESULT 60

T32735
 telomerase-associated protein component 1 - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
 C;Accession: T32735
 R;Nakayama, J.; Salto, M.; Nakamura, H.; Matsura, A.; Ishikawa, F.
 Cell 88, 875-884, 1997
 A;Title: TLP1: a gene encoding a protein component of mammalian telomerase is a novel m
 A;Reference number: Z21217; MUID:97236507; PMID:9118230
 A;Accession: T32735
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-2629 <NAP>
 A;Cross-references: EMBL:U89282; NID:91932816; PID:91932817; PIDN:AA851690.1
 A;Experimental source: strain Fischer 344
 C;Genetics:
 A;Gene: TLP1
 C;Function:
 A;Description: modification of telomerase-associated protein component 1 may regulate t

Query Match 8.0%; Score 105; DB 2; Length 2629;
 Best Local Similarity 21.7%; Pred. No. 23;
 Matches 62; Conservative 40; Mismatches 82; Indels 102; Gaps 14;

QY 9 HLLVLQIALLPATQGNKVLGKKDPVEL-----TCTASQKSI- 49
 Db 2235 HPLVLCQIHTLQHSQVPTAAASASGLLTSNDSVRLWQIPKEADPTCKPRSSAVIT 2294
 QY 50 QPHMKNSNQIKLIGNQGSFLT-----KGPSEKND-----RADSRSL 86
 Db 2295 AVAMAPDGSILVSGNAGELTLMQKAQAVATAPAPRVSDLWCSANAFVLISANENVSE 2354
 QY 87 WD-----QG---NPLIINKLIKIEDSDTYICEVEDQKEVOLVLFGLTANSDTHLQGS 138
 Db 2355 WQVELARKSGCTCNFRILYKRVLEDIG-----VLTMALAPD-----QGS 2394
 QY 139 LTL-----TLSPGSSPSVQCRSPRGKN-----IQGKTLVSQLELQDS 179
 Db 2395 LILMKEDVELQWKPFSTSSICRYAVHSSILCTSKDGLFYLQGSNGSGSILIEQES 2454
 QY 180 GTWCTVLONOKKVEFKIDI-VPRASALPAPPGSALPDQTSAL 224
 Db 2455 GKF-----EKLTDENLNLNPNNGS-----PVSTIQAEPESSSL 2488

RESULT 61

A32579
 neuroglian - fruit fly (Drosophila melanogaster)
 C;Species: Drosophila melanogaster
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: A32579

[illegible]

Db 409 LALAPTEENPMKKKILAAKGCVIIIECKPKAPKPKFSW-----SKGT 452

Qy 74 SKLNDRADSRSLWDGNFPLIIKNIKIEDSDPIICEVEQKEVQLLVFGLTANSDTHL 133

Db 453 EWLVN--SSRIILWEDGS--LEINNIIRNDGCIYTCAENNRK-----ANS----- 495

Qy 134 LOGGSILTLTESPP-----GSSPSVQCRS----- 157

Db 496 ---TGLVITDPFRILLAPINADITVGENATMQCAASFDPALDLTFWWSFNGYVIDFNK 551

Qy 158 -----PRGKNIQGGKTLVSQLELDQSGTWITCVLQNGKVEPFKIDIVPRASALPAPPTG 212

Db 552 ENIHQGNFPLDNGSELLRNQQLKHAGRYTCTAQITVDNSSASADLVVRGP--PGPG 609

Qy 213 SALPDPQTAS 222

Db 610 IRIEDIRATS 619

RESULT 66

S05944

neuronal cell surface protein F3 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 21-Jan-2000

C:Accession: S05944

A:Gene: Gibb11, G.; Rougon, G.; Mattei, M.G.; Goridis, C.

J: Cell Biol. 109, 775-788, 1989

A:Title: The mouse neuronal cell surface protein F3: a phosphatidylinositol-anchored mem

A:Reference number: S05944; MUID:89340657; PMID:2474555

A:Accession: S05944

A:Molecule type: mRNA

A:Residues: 1-1020 <SEN>

A:Cross-references: EMBL:X14943; NID:G50937; PIDN:CAA33075.1; PID:G50938

C:Genetics:

A:Map position: 15F

C:Superfamily: connecting, fibronectin type III repeat homology; immunoglobulin homology

F:1-20/Domain: signal sequence #status predicted <SIG>

F:256-312/Domain: immunoglobulin homology <IMM>

Query Match 7.9%; Score 104; DB 2; Length 1020;

Best local Similarity 20.2%; Pred. No. 8.8;

Matches 51; Conservative 28; Mismatches 89; Indels 84; Gaps 9;

Qy 16 LALPAPATQG--NKVVLGKKGDITVELCTASQKKSIGFHMKNNSQIKLNGSFLTKGP 73

Db 409 LALAPTEENPMKKKILAAKGCVIIIECKPKAPKPKFSW-----SKGT 452

Qy 74 SKLNDRADSRSLWDGNFPLIIKNIKIEDSDPIICEVEQKEVQLLVFGLTANSDTHL 133

Db 453 EWLVN--SSRIILWEDGS--LEINNIIRNDGCIYTCAENNRK-----ANS----- 495

Qy 134 LOGGSILTLTESPP-----GSSPSVQCRS----- 157

Db 496 ---TGLVITDPFRILLAPINADITVGENATMQCAASFDPALDLTFWWSFNGYVIDFNK 551

Qy 158 -----PRGKNIQGGKTLVSQLELDQSGTWITCVLQNGKVEPFKIDIVPRASALPAPPTG 210

Db 552 ENIHQGNFPLDNGSELLRNQQLKHAGRYTCTAQITVDNSSASADLVVRGP--PGPG 609

Qy 211 TGSALPDPQTAS 222

Db 610 GGLRIEDIRATS 621

RESULT 67

A57112

contactin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 21-Jan-2000

C:Accession: A57112

R:Peptides: E.; Nativ, M.; Campbell, P.L.; Sakurai, T.; Martinez, R.; Lev, S.; Clary, D.O.

Cell 82, 251-260, 1995

A:Title: The carbonic anhydrase domain of receptor tyrosine phosphatase beta is a funct

A:Reference number: A57112; MUID:95354206; PMID:7628014
 A:Accession: A57112
 A:Structure: preliminary; nucleic acid sequence not shown; not compared with conceptual tree
 A:Molecule type: mRNA
 A:Residues: 1-1021 <PELV>
 C:Superfamily: connectin; fibronectin type III repeat homology; immunoglobulin homology
 C:Keywords: membrane protein; phosphatidylinositol linkage
 F:256-312/Domain: immunoglobulin homology <IMM>

Query Match 7.9%; Score 104; DB 2; Length 1021;
 Best Local Similarity 20.2%; Pred. No. 8.8;
 Matches 51; Conservative 28; Mismatches 89; Indels 84; Gaps 9;

```

Oy 16 LALLPAATQ--NKVLLKGGDTVELTCTASQKSIQIHWKNSNQIKILNGSFLTKGP 73
Db 409 LALLPFTFEMNPMKKILAKGRVITCKPKAPKPKFSW-----SKGT 452
Oy 74 SKLNDRADSRSLMDQGNFPIIKLIKIEDSDTYICEVEDQKEVQLVFGILTANSDTHL 133
Db 453 EMLVW--SSRLIMEDGS--LEINNTIRNDGITYCFENNRK-----ANS----- 495
Oy 134 LQGSGLTITLESPP-----GSSPSVQGRS----- 157
Db 496 ---TGLVITNPFTIILAPINADITVGENATMQCAAFDPDLTFVWSFNGYVIDENK 551
Oy 158 -----PRGNIOGKRLSVSQLELDQSGTWTCTVLONOKKVEKIDIVPRASALPRP 210
Db 552 EITHIHVQRNFMLDANGELLIRNMQLKAGRYCTTAQTTVDNSSASADLVVRGP--FGPP 609
Oy 211 TGSALPDPQTAS 222
Db 610 GGLREDIRATS 621

```

RESULT 68
 S18731
 Ig kappa chain V-J region (MSI-N17) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 19-Mar-1997 #sequence_revision 05-Dec-1998 #text_change 23-Jul-1999
 C:Accession: S18731
 R:Hitama, T.; Takehita, S.; Yoshida, Y.; Yamagishi, H.
 Immunol. Lett. 27, 19-24, 1991
 A:Title: Structure of extrachromosomal circular DNA generated by immunoglobulin light c
 A:Reference number: S18731; MUID:91209891; PMID:1902191
 A:Accession: S18731
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-103 <HIR>
 A:Cross-references: EMBL:X54753; NID:955294; PION:CAA3855.1; PID:g55295
 A:Experimental source: spleen, strain BALB/c-nu/nu
 A>Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1990
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 18
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:11-85/Domain: immunoglobulin homology <IMM>
 F:18-83/Disulfide bonds: #status predicted

Query Match 7.9%; Score 103.5; DB 2; Length 103;
 Best Local Similarity 39.0%; Pred. No. 0.63;
 Matches 30; Conservative 5; Mismatches 37; Indels 5; Gaps 2;

```

Oy 34 GDTVELTCTASQKSIQIHWKNSNQIKILNGSFLTKGPSKLNDRADSRSLMDQ--NF 92
Db 11 GDIVTMTQASQGSINLNMFGQKP---GKAPKLLIYGASNLIEDGVPSRSGSRGTDF 66
Oy 93 PLIINKLIKIEDSDTYIC 109
Db 67 TLTISLEDEDMATYFC 83

```

RESULT 69
 KVM506

Ig kappa chain V region (J606) - mouse (tentative sequence)
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 31-Mar-2000
 C:Accession: A92811; A01929
 R:Johnson, N.; Stankard, J.; Paul, L.; Hood, L.
 J. Immunol. 128, 302-307, 1982
 A:Title: The complete V domain amino acid sequences of two myeloma inulin-binding prote
 A:Reference number: A92811; MUID:8209361; PMID:6798111
 A:Accession: A92811
 A:Molecule type: protein
 A:Residues: 1-108 <JCH>
 C:Comment: This chain was isolated from a myeloma protein that binds inulin.
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 1
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:16-90/Domain: immunoglobulin homology <IMM>
 F:23-88/Disulfide bonds: #status predicted

Query Match 7.9%; Score 103.5; DB 1; Length 108;
 Best Local Similarity 39.0%; Pred. No. 0.67;
 Matches 30; Conservative 5; Mismatches 37; Indels 5; Gaps 2;

```

Oy 34 GDTVELTCTASQKSIQIHWKNSNQIKILNGSFLTKGPSKLNDRADSRSLMDQ--NF 92
Db 16 GDIVTMTQASQGSINLNMFGQKP---GKAPKLLIYGASNLIEDGVPSRSGSRGTDF 71
Oy 93 PLIINKLIKIEDSDTYIC 109
Db 72 TLTISLEDEDMATYFC 88

```

RESULT 70
 S21668
 Ig kappa chain V region (24) - human
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
 C:Accession: S21668
 R:Rooschenthaler, F.; Schaeble, K.F.; Thiebe, R.; Zachau, H.G.
 Biol. Chem. Hoppe-Seyler 373, 177-186, 1992
 A:Title: Of orphans and UHOB. Delimitation of the germline repertoire of human immunogl
 A:Reference number: S21666; MUID:92281681; PMID:1596359
 A:Accession: S21668
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-117 <ROE>
 A:Cross-references: EMBL:X64642; NID:9432651; PIDN:CAA45916.1; PID:91335151
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:38-112/Domain: immunoglobulin homology <IMM>

Query Match 7.9%; Score 103.5; DB 2; Length 117;
 Best Local Similarity 33.0%; Pred. No. 0.74;
 Matches 37; Conservative 9; Mismatches 51; Indels 15; Gaps 4;

```

Oy 11 LVLVQLALLPAA-----TQANKVLLKGGDTVELTCTASQKSIQIHWKNSNQIKILGN 64
Db 9 LGLGLVLMPLGARDIQMTQSPSSLSASVDRVITTRASQGSNNLNMVQOKP---GK 64
Oy 65 QGSFLTKGPSKLNDRADSRSLMDQ--NFPIIKLIKIEDSDTYICEVED 113
Db 65 TPKFLIYAASLSQGSIRFS--DSGSGTDTLTISLQFEDPATYICQOQSD 114

```

RESULT 71
 KJHWK
 Ig kappa chain precursor V-1 region (Walker) - human
 C:Species: Homo sapiens (man)
 C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 21-Jan-2000
 C:Accession: A01883
 R:Klobeck, H.G.; Combratio, G.; Zachau, H.G.
 Nucleic Acids Res. 12, 6995-7006, 1984
 A:Title: Immunoglobulin genes of the kappa light chain type from two human lymphoid cel

A:Reference number: A93534; MUID:85014148; PMID:6091049
 A:Accession: A01883
 A:Molecule type: DNA
 A:Residues: 1-129 <KLO>
 A>Note: the sequence was determined from the differentiated gene
 C:Gene(s):
 A:Gene: GDB:IGKV1
 A:Cross-reference: GDB:136264
 A:Map position: 2p12-2p12
 C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IgH and IGM, the subunits associate into 16 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin
 F:1-22/Domain: signal sequence; #status predicted <SIG>
 F:23-129/Product: Ig kappa chain V-I region (Walker) #status predicted <MAT>
 F:23-45/Region: framework 1
 F:38-112/Domain: immunoglobulin homology <IMM>
 F:46-56/Region: complementarity-determining 1
 F:57-71/Region: framework 2
 F:72-78/Region: complementarity-determining 2
 F:79-110/Region: framework 3
 F:111-119/Region: complementarity-determining 3
 F:120-129/Region: framework 4
 F:45-110/Disulfide bonds: #status predicted

Query Match 7.9%; Score 103.5; DB 1; Length 129;
 Best Local Similarity 32.8%; Pred. No. 0.83;
 Matches 43; Conservative 10; Mismatches 61; Indels 17; Gaps 5;

Db 1 MNRGVPFRLHLLVQLALPAA-----TGQNKVVLGKKGDTVELTCTASOKSIOFHWK 54
 1 MDRKVPKQ--LIGILLMLKAGRCIDIQTOSPSLSVGRDRTITRAQSSINLIWY 58

QY 55 NSNOIKILNGSFLTKGPKLNDRADSRSLMDQG--NPLIIKLIKIEDSDTYICEVD 113

Db 59 QQKP-----GKAPKLLIYAASSLGCVTSRFSGSGTDFTLITSLQPEDSATVYC---- 110

QY 114 QKEVQLLVFG 124

Db 111 QQSXYTLITFG 121

RESULT 72

A53449
 Plasmacytoma-associated neuronal glycoprotein PANG - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 24-Sep-1999
 C:Accession: A53449
 R:Connolly, M.A.; Grady, R.C.; Mushinski, J.F.; Marcu, K.B.
 Proc. Natl. Acad. Sci. U.S.A. 91, 1337-1341, 1994
 A:Title: PANG, a gene encoding a neuronal glycoprotein, is ectopically activated by int
 A:Reference number: A53449; MUID:94151325; PMID:8108413
 A:Accession: A53449
 A:Status: preliminary
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1028 <CON>
 A:Cross-reference: GB:L01991; NID:g200056; PIDN:AAAI7403.1; PID:g200057
 C:Superfamily: connectin; fibronectin type III repeat homology; immunoglobulin homology
 C:Keywords: glycoprotein

Query Match 7.9%; Score 103.5; DB 2; Length 1028;
 Best Local Similarity 22.1%; Pred. No. 9.6;
 Matches 53; Conservative 33; Mismatches 83; Indels 71; Gaps 10;

QY 23 TQGNKVVLGKK-----GDTVELTCTASOKK--SIOFHWKNSNOIKILNGSFLTK-GP 73
 Db 497 TEPRIILASNMVAVGESVILPCQVQHPDLIDIMFAMVFNGLITPKDGSIFKAVG 556

QY 74 SKLNDRADSRSLMDQGNPLIIKLIKIEDSDTYICEVEDQKEVY-----QLLVFG----- 124

Db 557 SSSGD-----LMIRNIOLKISGKVCVQVGVSSAAELIVRGSPGP 601

QY 125 -----LTANSDTLLOGQSLTLLESPPSSPVQCRSPRKNIGGKT 168

Db 602 ENVKDEITDTTQASWTEGTDH-----SVISYAVQAKTFPVGMSVAT 648

QY 169 LSVQLELODQSGTWTCTVLOKQKVEFKDIPRASALPAPPTGSLPDP--QTASALPD 226
 Db 649 VP-----EVIDGKTHATVVELNPNWVEYFRIVASNKIGGEP---SLPSEKRTTEBADE 701

RESULT 73

T20992
 hypothetical protein F15G9.4a - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T20992; T24733
 R:Subleat, J.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: Z19355
 A:Accession: T20992
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-5175 <W1>
 A:Cross-reference: EMBL:247068; PIDN:CAA87335.1; GSPDB:GN00028; CESP:F15G9.4a
 A:Experimental source: Clone F15G9
 R:Kershaw, J.
 submitted to the EMBL Data Library, December 1994
 A:Reference number: Z19929
 A:Accession: T24733
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-5175 <W12>
 A:Cross-reference: EMBL:247070; PIDN:CAA87344.1; GSPDB:GN00028; CESP:F15G9.4a
 C:Experimental source: clone T09B9
 C:Gene(s):
 A:Gene: CESP:F15G9.4a
 A:Map position: X
 A:Introns: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3;
 ; 2512/2; 2593/3; 2699/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1;
 1; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4838/1; 4878/1; 4941/1; 5011/1; 5077/1;

Query Match 7.9%; Score 103.5; DB 2; Length 5175;
 Best Local Similarity 18.7%; Pred. No. 65;
 Matches 62; Conservative 50; Mismatches 108; Indels 111; Gaps 16;

QY 26 NKVVLGKKGDTVELTCTASOKSIOFHW-----KSNQIKILG-----NQ 65
 Db 709 NKQWVG-RGDRVSEFCKTIRGPHPIRPFKNGKDIKPDYIKINEGOLHWGABEDA 767

QY 66 GSF-----LTGPKSLNDRADSRSL-----MD 88
 Db 768 GAVSCGEMMAGKDVAVANLSVGRVPTTIESHTVAVNIERQVTLQCLAVGIRPPEIENG 827

QY 89 QGNF-----PLIIKLIKIEDSDTYICEVD---QKEVQLLVFGTLA- 127
 Db 828 KGNVLLATINPNRYTQLADGNLLITDAQIEDQGFCTIRANRYGQOSGSTLLMVTGLVSP 887

QY 128 -----NSDTHLQGGSLTLT-----LESPGSSPVQCRSP--RGKIQO---GGKTLVSQ 173
 Db 888 VLGHPPEBQLIEGQDLTSLSCVVLGTFRPSVWIMKDPVBEGRPTIKIEGGSLRLRG 947

QY 174 LELQDSGTWTCVTLQ--NQKVEFKDIPRASALPAPPTGSLPDPQTASALPDPPAAS 231
 Db 948 GNRKDGKKTCTCLAVASNGSTLHINQLIKKEFPVKKPGGIVF-----KPTIS 996

QY 232 ALPAA-LAVIS-----FLGLGLGVACVLTART 257
 Db 997 GMDKHAVALVNSTHVDLDEGFAIPCVVSGT 1027

RESULT 74

T43290
 hemiscentin precursor - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 18-Feb-2000

C/Accession: T43290; T20993; T24734
 R/Vogel, B.E.; Hedgecock, E.M.
 Submitted to the EMBL Data Library, June 1998
 A/Description: Hemiscentin is required for hemidesmosome mediated cell adhesion and germ
 A/Reference number: 222396
 A/Accession: T43290
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-5198 <VOG>
 A/Cross-references: EMBL:AF074901; PIDN:AA026792.1
 R/Sulston, J.
 Submitted to the EMBL Data Library, December 1994
 A/Reference number: 219355
 A/Accession: T20993
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-5198 <ML>
 A/Cross-references: EMBL:Z47068; PIDN:CAA87336.1; GSPDB:GN00028; CESP:FL5G9.4b
 R/Kershaw, J.
 Submitted to the EMBL Data Library, December 1994
 A/Reference number: 219929
 A/Accession: T24734
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-5198 <WT>
 A/Cross-references: EMBL:Z47070; PIDN:CAA87345.1; GSPDB:GN00028; CESP:FL5G9.4b
 A/Experimental source: clone T0989
 C/Genetic8:
 A/Gene: hlm-4; FL5G9.4b
 A/Map position: X
 A/Intons: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3;
 12512/2; 2593/3; 2699/3; 2759/1; 2852/1; 2889/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1;
 1; 4225/1; 4361/1; 4408/1; 4436/1; 4458/1; 4647/3; 4838/1; 4902/1; 4964/1; 5034/1; 5100/1
 Query Match 7.9%; Score 103.5; DB 2; Length 5198;
 Best Local Similarity 18.7%; Pred. No. 65;
 Matches 62; Conservative 50; Mismatches 108; Indels 111; Gaps 16;
 Db 26 NKVVLGKKGDTVELTC-TASQKKSIOFHW-----KNSNQIKILG---NQ 65
 Db 709 NKMMVG-RGDRVSPECKTIRGKPHKIRFWKNGKDLKPDYIKINEGQHLIMGAKDEDA 767
 Db 66 GSF-----LTGPKSLNDRADSRSL-----WD 88
 Db 768 GAYSCVGENMAGKDVQVANLSVGRVPTTIESPHTVRVNIERQVTLQCLANGIPPEIEMQ 827
 Db 89 QGNF-----PLIINKLKIEDSDTYICEVED---OKEEVQLVFGTLA- 127
 Db 828 KGNVLATLNNPRYTQQLADGNLLITPDQIEDQGFCTARNTYGOQSQSTLMTGLVSP 887
 Db 128 ----NSDTHLLQGQSITLT---LESPPGSSPSVQCRSP--RGNKIQ--GKTLVSQ 173
 Db 888 VLGVHPPEEQLEIEGQDLTLCVVVLGTPKPSIWKDKDPVEBPPTIKIEGGSLLRLRG 947
 Db 174 LELODSGWTCTVQ--NQKVERKIDIVPRASALPAPPTGSALPDQTASALDPDPAAS 231
 Db 948 GNPDEGKYTCIAVSPAGNSTLHINVOLIKKPEFVYKREGGIVF-----KPTIS 996
 Db 232 ALPAA-LAVIS-----FLTGLGLGVAACVLART 257
 Db 997 GMDKHYAVVNSTHVDLDGEGFALPCVSGT 1027

RESULT 75
 A44783
 ecto-ATPase precursor - rat
 N/Alternate names: canalicular bile acid transport protein
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 23-Jul-1999
 C/Accession: A44783; A44410
 R/Lin, S.H.; Guidotti, G.
 J. Biol. Chem. 264, 14408-14414, 1989

A/Title: Cloning and expression of a cDNA coding for a rat liver plasma membrane ecto-A
 A/Reference number: A44783; MUID:9340561; PMID:2527235
 A/Accession: A44783
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-519 <LIN>
 A/Cross-references: GB:04963; NID:9203969; PIDN:AAA41104.1; PID:9203990
 R/Sippel, C.J.; Suchy, F.J.; Nanthanarayanan, M.; Perlmuter, D.H.
 J. Biol. Chem. 268, 2083-2091, 1993
 A/Title: The rat liver ecto-ATPase is also a canalicular bile acid transport protein.
 A/Reference number: A44410; MUID:9313196; PMID:8420979
 A/Accession: A44410
 A/Molecule type: Protein
 A/Residues: 110-120;122-138;148-150 <SIP>
 C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
 C/Keywords: glycoprotein; liver; phosphoprotein; transmembrane protein
 F/1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
 F/160-217/Domain: immunoglobulin homology <IMM1>
 F/252-301/Domain: immunoglobulin homology <IMM2>
 F/337-394/Domain: immunoglobulin homology <IMM3>
 Query Match 7.8%; Score 103; DB 2; Length 519;
 Best Local Similarity 20.9%; Pred. No. 47;
 Matches 51; Conservative 34; Mismatches 93; Indels 66; Gaps 7;
 Db 33 KGDVELTCASQKKSIOFHWKNSNQIKILGNGSFLTKPSKLNDRADSRSLMDQGNF 92
 Db 251 QGSNLTNSCHADSNPPAQYFWLINELKQTSQSE----- 283
 Db 93 PLIINKLKIEDSDTYICEVEDOKEEVQLVFGTLAN-----SDTHLL 134
 Db 284 -LFTSNTTNGSGTYACFVNN-----TVTGLSRTYKNIITVEPVTQPSIQITNTTVK 335
 Db 135 QGQSITLTLESPPG-----SSPSVQCSPPGKNIQGGKTLVSQLELODSGTWTCTV 186
 Db 336 ELGSVTLTCSKDTGVSVRWLFNSQSLQ-TDWMTLSDQNSTLRIDPIKREDAGDYQCEI 394
 Db 187 LQNGKVEKIDIVPRASALPAPPTG-SALPDQTSALDPDPAASALPALAVISFLG 245
 Db 395 ---SNPFSRISHPIKIDVLPDPTQNGSLSEGALIGIVGSVAGVALPALAVLSRK 451
 Db 246 LGIG 249
 Db 452 TGGG 455

RESULT 76
 149583
 differentiation antigen - mouse
 A/Species: Mus musculus (house mouse)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
 C/Accession: 149583
 R/Law, C.
 J. Immunol. 151, 175-187, 1993
 A/Title: Organization of the murine Cd22 locus. Mapping to chromosome 7 and characteriz
 A/Reference number: 149583; MUID:93315834; PMID:8100643
 A/Accession: 149583
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-862 <RES>
 A/Cross-references: GB:146928; NID:9348965; PIDN:AAA02562.1; PID:9348966
 C/Genetic8:
 A/Gene: CD22
 Query Match 7.8%; Score 102.5; DB 2; Length 862;
 Best Local Similarity 21.9%; Pred. No. 9.2;
 Matches 52; Conservative 33; Mismatches 85; Indels 67; Gaps 10;
 Db 32 KGDVELTC--TASQKKSIOFHWKNSNQIKILGNGSFLTKPSKLNDRADSRSLMDQ 89
 Db 535 RAGQRYVLQCDFAESNPAEVRFFWKNGSLVQGRYLSFGSVSP-----DS 581
 Db 90 GNPFLIINKLKIEDSDTYICEVEDOKEEVQLV-----FGLTANSDFHLGQGSITLTLS 145

```

Db      582 GATNCNMNN-----SIEETLSQAMNLOVLAAPRRRLVSIIPGHVEGKATLSCS 633
QY      146 ---PPGSSPSVQCRSPRGKNI-OQGKTLVSQLELODSGTWTCTVLONOKRVEKDIYV 201
Db      634 DAPPIISQYWPDS--GQDLHSSGQKLRLEPLEVQHTGSYRC-----674

QY      202 RASALPAPPTGSALEPQTASALPDPAPASALPAALAVISFLGLGIG--VACVLA 255
Db      675 --KGTNGIGTGESPPSTLTLYYSPE-----TIGKVALGIGFCLTITCILA 717

RESULT 77
S51556
Vascular endothelial growth factor receptor 1 precursor - Japanese quail
N:Alternate names: quail endothelial kinase 1 (Quek 1); vascular endothelial growth fac
C:Species: Coturnix coturnix japonica (Japanese quail)
C>Date: 07-May-1995 #sequence revision 01-Sep-1995 #text_change 08-Dec-2000
C:Accession: J04953; A56598; T51162; S51656
R:Etchmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.
Gene 174, 3-8, 1996
A:Title: Molecular cloning of Quek 1 and 2, two quail vascular endothelial growth factor
A:Reference number: J04953; MUID:97017121; PMID:8863722
A:Accession: J04953
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1348 <EIC2>
A:Cross-references: EMBL:X83288; NID:G603523; PIDN:CAA58268.1; PID:G603524
A:Note: submitted to the EMBL Data Library, December 1994
R:Etchmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.
Mech. Dev. 42, 33-48, 1993
A:Title: Two molecules related to the VEGF receptor are expressed in early endothelial c
A:Reference number: A56598; MUID:93378866; PMID:8336413
A:Contents: E16 spinal cord
A:Accession: A56598
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 310-1348 <EIC2>
A:Cross-references: GB:S65205; NID:G410680; PIDN:AA828127.1; PID:G410681
A:Note: sequence extracted from NCBI backbone (NCBIN:137162, NCBIPI:137163)
R:Marcelle, C.; Etchmann, A.
Oncogene 7, 2479-2487, 1992
A:Title: Molecular cloning of a family of protein kinase genes expressed in the avian en
A:Reference number: I50595; MUID:93096482; PMID:1281306
A:Accession: I51162
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1023-1079 <MAR>
A:Cross-references: EMBL:X69694; NID:G395226; PIDN:CAA49364.1; PID:G938278
A:Note: the species is not identified by the authors; the most probable species is shown
C:Comment: This protein is an endothelial-specific receptor and binds vascular endothel
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; embryo; growth factor receptor; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:120-148/Product: vascular endothelial growth factor receptor 1 #status predicted <MAR>
F:756-777/Domain: transmembrane #status predicted <TM>
F:823-1160/Domain: protein kinase homology <KIN>
F:831-839/Region: protein kinase ATP-binding motif

Query Match      7.8%; Score 102.5; DB 2; Length 1348;
Best Local Similarity 22.9%; Pred. No. 16;
Matches 49; Conservative 33; Mismatches 73; Indels 59; Gaps 12;

QY      9 HLLVLIQALLPAATQGNKVVLLGKGDVETCTASQKSIQFHW-KNSNQIKLNGQS 67
Db      536 HVTGRLINLQPRSQLTEK-----DNTSLQCTADKFTFKLSWYKLSYTHV---SQT 584
QY      68 FLITG--PSKINDAADSRL-----MDQGFPLIKLKEDSDTYICEVDQKE 117
Db      585 F--GGLPMPYCKNLDAQLQKLNATVSNVNGENTVLEILIRISLQDGDYVCAIQDKAK 641
QY      118 VQ-LLVGLTANSDTH-----LLQGSLTTLLESPSPSSPSVQCRSPRG-----KN 162

```

```

Db      642 TQHCIVKHLTVQEPRLPRLVGLNQTNI-----GRIEVLG-C-TVNGVPPENITWPKN 694
QY      163 IQ-----GKTLVSQLELODSGTWTC 184
Db      695 SETLPEDSGIVLKDGNKTLTIRVRKEDGLTYTC 728

RESULT 78
WMM5R1
Biliary glycoprotein A precursor - mouse
N:Alternate names: carcinoembryonic antigen mncGM1a; murine hepatitis virus receptor
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text_change 22-Jun-1999
C:Accession: J01505; A49006; A41563; S11625; S11625
R:McCaig, K.; Rosenberg, M.; Nedellec, P.; Turbide, C.; Beauchemin, N.
Gene 127, 173-183, 1993
A:Title: Expression of the Bgp gene and characterization of mouse colon biliary glycopro
A:Reference number: J01505; MUID:93273228; PMID:8500759
A:Accession: J01505
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-458 <DVE>
A:Cross-references: GB:M77196; NID:G194002; PIDN:AAA37858.1; PID:G194003
R:Beauchemin, N.
Submitted to the EMBL Data Library, May 1990
A:Reference number: S11626
A:Accession: S11626
A:Molecule type: mRNA
A:Residues: PQ, 82-458 <BEA>
A:Cross-references: EMBL:X15351
R:Beauchemin, N.; Turbide, C.; Afar, D.; Bell, J.; Raymond, M.; Stanners, C.P.; Fuks, A
Cancer Res. 49, 2017-2021, 1989
A:Title: A mouse analogue of the human carcinoembryonic antigen.
A:Reference number: S11625; MUID:89195121; PMID:2702644
A:Accession: S11625
A:Molecule type: mRNA
A:Residues: PQ, 82-321 <BE2>
A:Cross-references: EMBL:X15351
C:Comment: This protein is expressed at the cell surface and plays a determinant role in
C:Genetics:
A:Gene: Bgpa
C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
C:Keywords: glycoprotein; receptor; transmembrane protein
F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-458/Product: biliary glycoprotein A #status predicted <MHR>
F:160-219/Domain: immunoglobulin homology <IMM1>
F:254-303/Domain: immunoglobulin homology <IMM2>
F:339-396/Domain: immunoglobulin homology <IMM3>
F:429-447/Domain: transmembrane #status predicted <TMN>
F:449-458/Domain: intracellular #status predicted <CYT>
F:71,89,104,148,152,199,206,210,226,258,290,294,304,317,333,375/Binding site: carbohydrate
F:167-217,261-301,346-394/Disulfide bonds: #status predicted

Query Match      7.7%; Score 102; DB 1; Length 458;
Best Local Similarity 22.5%; Pred. No. 4;
Matches 42; Conservative 35; Mismatches 68; Indels 42; Gaps 10;

```

Oy 16 LALLPAATQGNKVVLGKKDTEVELTCTA-SOKKSIQFHMKNSNOIKILGNGSFLTKGPS 74
 Db 143 ILKRNITNSNPNV-BGDDSVSLTCDSTYDPDNTNVMYSRN-----GESLSEGD- 191
 Oy 75 KLNDRAHSRSLMDQGNFPLIIRKLIKEDSDTYICEVE-----DOKEEVOL-LVEFG---- 124
 Db 192 -----RLKLSBGNNLTLLLVNTNRNDTPGYCETNPVSVNRSDPFLNITTYGPTP 242
 Oy 125 LTANSDFHLLOGSLTLTLESPPGSSPEVQC-----RSPRKNIQGGKTLSSVSOLELQDS 179
 Db 243 IISPSDIYLRHPSGNLNSCHA--ASNPPAQYFWLINERPHASS-----QELFIPMITTNS 296
 Oy 180 GTWTCTV 186
 Db 297 GTYTCFV 303

RESULT 79

cell adhesion molecule short form (cell-CAM105) - rat
 N/Alternate names: C-CAM protein
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 22-Nov-1993 #sequence revision 01-Sep-1995 #text_change 23-Jul-1999
 C/Accession: S23969; S32483; S38826; S10563; S32102
 R/Critic: O.; Huang, O.H.; Flanagan, D.; Hixson, D.; Lin, S.H.
 Biochem. J. 285, 47-53, 1992
 A/Title: Molecular cloning and expression of a new rat liver cell-CAM105 isoform. Differ.
 A/Reference number: S23969; PMID:93344597; PMID:1637321
 A/Accession: S23969
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-458 <CUI>
 A/Cross-references: EMBL:Z12019; NID:G55858; PIDN:CAA78054.1; PID:G55859
 R/Edlund, M.; Gaardsvoll, H.; Bock, E.; Oebdrink, B.
 Eur. J. Biochem. 213, 1109-1116, 1993
 A/Title: Different isoforms and stock-specific variants of the cell adhesion molecule C-
 A/Reference number: S32483; PMID:93279310; PMID:8504806
 A/Accession: S32483
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-458 <EDL>
 A/Cross-references: EMBL:X71122; NID:G287819; PIDN:CAA50435.1; PID:G287820
 R/Cheung, P.H.; Cullis, O.; Otu, Y.; Batley, K.; Thompson, N.; Hixson, D.C.; Lin, S.H.
 Biochem. J. 293, 427-435, 1993
 A/Title: The cytoplasmic domain of C-CAM is required for C-CAM-mediated adhesion function.
 A/Reference number: S38826; PMID:94058980; PMID:8240240
 A/Accession: S38826
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-48, 'A', 50-54, 'A', 56-69, 'G', 71-72, 'LNPV', 77-85, 'D', 87, 'M', 89, 'K', 91, 'G', 93-
 A/Cross-references: EMBL:M92848; NID:G20366; PIDN:AAA16783.1; PID:G20367
 R/Autivillius, M.; Hansen, O.C.; Lazrek, M.B.S.; Bock, E.; Oebdrink, B.
 FEBS Lett. 264, 267-269, 1990
 A/Title: The cell adhesion molecule cell-CAM 105 is an ecto-ATPase and a member of the I
 A/Reference number: S10563; PMID:90292222; PMID:2141577
 A/Accession: S10563
 A/Molecule type: protein
 A/Residues: 'X', 58-66, 'A', 68, 121-124, 'F', 126, 'Q', 128-134, 'X', 136-138, 'X', 1356-160, 'X', 1362
 C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
 F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
 F:160-217/Domain: immunoglobulin homology <IMM1>
 F:252-301/Domain: immunoglobulin homology <IMM2>
 F:337-394/Domain: immunoglobulin homology <IMM3>
 Query Match 7.7%; Score 102; DB 2; Length 458;
 Best Local Similarity 21.0%; Pred. No. 4.8;
 Matches 49; Conservative 33; Mismatches 85; Indels 66; Gaps 7;
 Oy 33 KGDTEVELTCTASQKKSIOFHMKNSNOIKILGNGSFLTKGPSKLNDRAHSRSLMDQGNF 92
 Db 251 QGSNLTNSCHADSNPPAQYFWLINERKLTSSQE----- 283

Oy 93 PLIKNKIEDSDTYICEVEDQKEVOLVFGLTAN-----SDTHLL 134
 Db 284 -LFLSNITTNNSGYACFVNN-----TYTGLSRTTVKNIITVEPTQSIQITNTTVK 335
 Oy 135 QGOSLTLTLESPPG-----SSPVQCRSPRKNIQGGKTLSSVSOLELQDSGTWCTV 186
 Db 336 ELGSLVTLCFSKOTGVSVRLFNLSQSLQ--TDHWTLSQDNTLIDIPKREDADYQCEI 394
 Oy 187 LQNKQVEFKIDIVPRASALPAPPTG-SALPDQPTASALPDPAPASALPALA 238
 Db 395 ---SNPVSRISHPIKIDVIPDPQTGNSGLSEGAIAGIVIGSVAGVALLA 444

RESULT 80

C-CAM2a protein isoform precursor - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 15-Feb-1997 #sequence revision 13-Mar-1997 #text_change 20-Jun-2000
 C/Accession: S68177
 R/Lucka, L.; Cichocka, I.; Baeumler, K.; Bechler, K.; Reutter, W.
 Eur. J. Biochem. 234, 527-535, 1995
 A/Title: A short isoform of carcinoembryonic-antigen-related rat liver cell-cell adhesi
 A/Reference number: S68177; PMID:96128184; PMID:8536699
 A/Accession: S68177
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-458 <LUC>
 A/Cross-references: EMBL:X91137; NID:G160272; PIDN:CAA62577.1; PID:G160273
 C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
 F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEA>
 F:133/Domain: signal sequence #status predicted <SIG>
 F:34-458/Product: C-CAM2a protein isoform #status predicted <MAT>
 F:252-301/Domain: immunoglobulin homology <IMM>
 Query Match 7.7%; Score 102; DB 2; Length 458;
 Best Local Similarity 21.0%; Pred. No. 4.8;
 Matches 49; Conservative 33; Mismatches 85; Indels 66; Gaps 7;
 Oy 33 KGDTEVELTCTASQKKSIOFHMKNSNOIKILGNGSFLTKGPSKLNDRAHSRSLMDQGNF 92
 Db 251 QGSNLTNSCHADSNPPAQYFWLINERKLTSSQE----- 283
 Oy 93 PLIKNKIEDSDTYICEVEDQKEVOLVFGLTAN-----SDTHLL 134
 Db 284 -LFLSNITTNNSGYACFVNN-----TYTGLSRTTVKNIITVEPTQSIQITNTTVK 335
 Oy 135 QGOSLTLTLESPPG-----SSPVQCRSPRKNIQGGKTLSSVSOLELQDSGTWCTV 186
 Db 336 ELGSLVTLCFSKOTGVSVRLFNLSQSLQ--TDHWTLSQDNTLIDIPKREDADYQCEI 394
 Oy 187 LQNKQVEFKIDIVPRASALPAPPTG-SALPDQPTASALPDPAPASALPALA 238
 Db 395 ---SNPVSRISHPIKIDVIPDPQTGNSGLSEGAIAGIVIGSVAGVALLA 444
 RESULT 81
 JCI508
 biliary glycoprotein D - mouse
 N/Alternate names: biliary glycoprotein 1
 C/Species: Mus musculus (house mouse)
 C/Date: 24-Feb-1994 #sequence revision 24-Feb-1994 #text_change 23-Jul-1999
 C/Accession: JCI508; S65940; S36852
 R/McCaig, K.; Rosenberg, W.; Nedellec, P.; Turbide, C.; Beauchemin, N.
 Gene 127, 173-183, 1993
 A/Title: Expression of the Bgp gene and characterization of mouse colon biliary glycopr
 A/Reference number: JCI505; PMID:93273228; PMID:8500759
 A/Accession: JCI508
 A/Molecule type: mRNA
 A/Residues: 1-521 <MCC>
 A/Cross-references: EMBL:X67279; NID:G50170; PIDN:CAA7696.1; PID:G50171
 A/Experimental source: strain CD1, tissue colon
 R/Nedellec, P.; Turbide, C.; Beauchemin, N.
 Eur. J. Biochem. 231, 104-114, 1995

A>Title: Characterization and transcriptional activity of the mouse biliary glycoprotein
 A/Reference number: S65939; MUID:95354678; PMID:7628460
 A/Accession: S65940
 A/Status: translation not shown
 A/Molecule type: DNA
 A/Residues: 1-21 <NED>
 A/Cross-references: EMBL:X64054; NID:g1039337
 A/Experimental source: strain BALB/c
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
 A/Comment: This protein is expressed at the cell surface and plays a determinant role in C/Genetics:
 A/Gene: Bgpl
 A/Map position: 7
 C/Superfamily: carcinoembryonic antigen, carcinoembryonic antigen precursor amino-terminal
 C/Keywords: glycoprotein; receptor; transmembrane protein
 F:1138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
 F:160-219/Domain: immunoglobulin homology <IMM1>
 F:254-303/Domain: immunoglobulin homology <IMM2>
 F:339-396/Domain: immunoglobulin homology <IMM3>
 F:71,89,104,148,199,206,210,226,258,290,294,304,317,333,375/Binding site: carbohydrate

Query Match 7.7%; Score 102; DB 2; Length 521;
 Best local similarity 22.5%; Pred. No. 5.5;
 Matches 42; Conservative 35; Mismatches 68; Indels 42; Gaps 10;

QY 16 LALLPAATGKGVKGLKKGTVELTCTA-SQKSIQPHMKNSQIKLNGSFLTKGPS 74
 DB 143 ILKPNITNSNPNV-EGDSDVSLTCDSTYDPDININLMSRN-----GESLSRGD- 191
 QY 75 KLNDRADRSRLMDGFPILIKKLTEDSDTYICEVE-----DOKEVOL-LVFG--- 124
 DB 192 -----RLKSEGRRTLTLLNVTNRNDGPVCEFRNVSVNRSDFSLNITIGPDP 242

QY 125 LTRANSDLHOGSLTLTLSPSPGSSPSVOC-----RSPKGIQGGKTLVSQLEILOS 179
 DB 243 IISPDYILHPSGLNLSCHA--ASNPAQYFMLEKPHASS-----QELFINITNNS 296

QY 180 GTWTCIV 186
 DB 297 GTYTCFV 303

RESULT 82
 138346
 elastic titin - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
 C/Accession: 138346
 R/Label: S.; Kolmerer, B.
 Science 270, 293-296, 1995
 A/Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
 A/Reference number: A57430; MUID:96026330; PMID:7569978
 C/Accession: 138346
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-7962 <RES>
 A/Cross-references: EMBL:X90569; NID:g1017426; PIDN:CAA62189.1; PID:g1017427
 C/Genetics:
 A/Gene: GDB:TTN
 A/Cross-references: GDB:127867; OMIM:188840
 A/Map position: 2q31-2q31

Query Match 7.7%; Score 102; DB 2; Length 7962;
 Best local similarity 23.6%; Pred. No. 1.4e+02;
 Matches 59; Conservative 22; Mismatches 81; Indels 88; Gaps 11;

QY 11 LTVGLALLPAATGKGVKGLKKGTVELTCTASQKSIQPHW-----KNSQIKILG 63
 DB 4245 LSVLEPHTYVEKPSIKVT--TSDTCTLECTVAGTPELSTKPKGKELTSDKTKYI-- 4299
 QY 64 NOGSFLTKGPSKLNDRADRSRLMDGFPILIKKLTEDSDTYICEVEDOKEVOLIVF 123

DB 4300 ----SFPNKVSG-----LKINVPASDGVSPFVQNP----- 4328
 QY 124 GLTNSDTHLLOGSLTLTLSP-----GSSPVQCR--SP-----R 159
 DB 4329 -VGKDSCTASLQVSDRTV---PFSFTRKLKETNGLSGSSVMECKVSPPIYSWFHE 4383
 QY 160 GKNIGGK-----TLVSQLEILOSGLTWTCITVQNGKVEFKIDIVRASALPA 208
 DB 4384 GNEISSGRKRYQTTLTDNLTALTNMLESDSGDITCIATNMAGSDSCAPLYR-----E 4438

QY 209 PPTGSALPDP 218
 DB 4439 PPSFVGKPPD 4448

RESULT 83
 PH1066
 Ig light chain V region (clone 165.14) - mouse (fragment)
 C/Species: Mus musculus (house mouse)
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
 C/Accession: PH1066
 R/Tilman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
 J. Exp. Med. 176, 761-779, 1992
 A/Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B
 A/Reference number: PH0971; MUID:92381444; PMID:1512540
 A/Accession: PH1066
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-98 <TIL>
 A/Experimental source: B cell, strain [NZB x NZW]F1
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: immunoglobulin
 F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 7.7%; Score 101.5; DB 2; Length 98;
 Best local similarity 31.8%; Pred. No. 0.83;
 Matches 35; Conservative 7; Mismatches 39; Indels 29; Gaps 3;

QY 34 GDTVELTCTASQKSIQPHMKNSQIKLNGSFLTKGPSKLNDRADRSRLMDG-NF 92
 DB 16 GDTITTCRASQININLMSYQKP-----GNIPKLIYKASNLHTGVPSPFGSGSGTDF 71

QY 93 PLIKKLTEDSDTYICEVEDOKEVOLIVFGLTANSDLHOGSLTLT 142
 DB 72 TLTISSLOPEDATATTC-----LQGSYPLT 97

RESULT 84
 IUMSNG
 neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse
 N/Alternate names: NCAM-120
 C/Species: Mus musculus (house mouse)
 C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2000
 C/Accession: A29673; S00382; A44290
 R/Bartelme, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Fontec
 EMO J. 6, 907-914, 1987
 A/Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000
 A/Reference number: A29673; MUID:87246524; PMID:3595563
 A/Accession: A29673
 A/Molecule type: mRNA
 A/Residues: 1-725 <PAR>
 A/Cross-references: EMBL:Y00051; NID:g53342; PIDN:CAA6263.1; PID:g53343
 R/Barbas, J.A.; Chaix, J.C.; Steinmetz, M.; Goridis, C.
 EMO J. 7, 625-632, 1988
 A/Title: Differential splicing and alternative polyadenylation generates distinct NCAM
 A/Reference number: S00382; MUID:88283628; PMID:3396534
 A/Accession: S00382
 A/Molecule type: DNA
 A/Residues: 642-656, 'D', 658-725 <BA2>
 A/Cross-references: EMBL:X07195
 R/Rougon, G.; Marshak, D.R.
 J. Biol. Chem. 261, 3396-3401, 1986
 A/Title: Structural and immunological characterization of the amino-terminal domain of "

A:Reference number: A44290; MUID:86140120; PMID:3512556

A:Accession: A44290

A:Molecule type: protein

A:Residues: 20-36 <R0U>

C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM molecule. Several forms of NCAM are produced by alternative splicing. See also PIR:J00688

C:Gene: NCAM

A:Gene: NCAM

A:Map position: 9

A:Introns: 701/1

C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin domain; cell adhesion; duplication; heparin binding; membrane

F:1-19/Domain: signal sequence #status predicted <SIG>

F:134-98/Domain: immunoglobulin homology <IMW1>

F:132-191/Domain: immunoglobulin homology <IMW2>

F:152-156/Region: heparin binding #status predicted

F:161-165/Region: heparin binding #status predicted

F:228-290/Domain: immunoglobulin homology <IMW3>

F:263-272/Region: NCAM binding #status predicted

F:323-388/Domain: immunoglobulin homology <IMW4>

F:420-482/Domain: immunoglobulin homology <IMW5>

F:519-596/Domain: fibronectin type III repeat homology <FN3A>

F:625-685/Domain: fibronectin type III repeat homology <FN3B>

F:141-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted

F:222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 101.5; DB 1; Length 725;

Best Local Similarity 19.9%; Pred. NO. 8.9;

Matches 46; Conservative 42; Mismatches 94; Indels 49; Gaps 10;

QY 3 RG-VEPRHLVLVQLALLPAPATQGNKVVIGKKGDVELTTPASQKSIQFHW-KNSNOK 60

DB 196 RGEINFKIQIVVNPPTVQARQSIYNATAMIGQSVTLVCADGEPPEPTMWTGDEPIE 255

QY 61 ILGNQSGFLTKGPKSLNDRAD--SPRSLMDQNPFLIKLKIENSDTYICEV---DQ 114

DB 256 -----NEEDERSRVSVDSE--VTRNDVKDDEAYCTIANKKAGEQ 297

QY 115 KEEOVLVFEV-----LTASDPTHLQGSQSLTLESPPSSPSVOCR-----SPRKNIQ 164

DB 298 DASILHKYPAKRIKYVENQTMELBEQVTLTCEASGPISITMRTSTRNISSEEDLD 357

QY 165 GG-----KTLVSQLELQDSGTWTC---TVLQNKKEVEFKIDIVP 202

DB 358 GIMVVRSHARVSSLTKSIOYRADAGYWCYCTASNTIGQSQSIDLEFYAPK 408

RESULT 85

IDENTIFICATION

neural cell adhesion molecule short domain form precursor - rat

N:Alternate names: NCAM-140

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 22-Jun-1999

C:Accession: S00846; B37795; I58136

R:Small, S.J.; Shull, G.E.; Santoni, M.J.; Akesson, R.

J. Cell Biol. 105, 2335-2345, 1987

A>Title: Identification of a cDNA clone that contains the complete coding sequence for a

A:Reference number: S00846; MUID:88059265; PMID:3680385

A:Accession: S00846

A:Molecule type: mRNA

A:Residues: 1-858 <SMA>

A:Cross-references: EMBL:X06564

R:Small, S.J.; Akesson, R.

J. Cell Biol. 111, 2089-2096, 1990

A>Title: Expression of the unique NCAM VASE exon is independently regulated in distinct

A:Reference number: A37795; MUID:91035620; PMID:1699951

A:Accession: B37795

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 340-381 <SM2>

R:Small, S.J.; Haines, S.L.; Akesson, R.A.

Neuron 1, 1007-1011, 1988

A>Title: Polypeptide variation in an N-CAM extracellular immunoglobulin-like fold is dev

A:Reference number: 158136; MUID:90166485; PMID:2483093
A:Accession: 158136
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 355-364 <RES>
A:Cross-references: GB:M2611; NID:G205643; PIND:AAA41679.1; PID:G205644
C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM molecule
C:Comment: Various forms of NCAM are produced by alternative splicing.
C:Genetics:
A:Gene: NCAM
C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin domain
C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; signal sequence
F:11-19/Domain: signal sequence #status predicted <SIG>
F:20-858/Product: neural cell adhesion molecule, short domain form #status predicted <M>
F:120-721/Domain: extracellular #status predicted <EXT>
F:34-98/Domain: immunoglobulin homology <IM1>
F:132-191/Domain: immunoglobulin homology <IM2>
F:152-166/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-290/Domain: immunoglobulin homology <IM3>
F:263-272/Region: NCAM binding #status predicted
F:323-398/Domain: immunoglobulin homology <IM4>
F:430-492/Domain: immunoglobulin homology <IM5>
F:529-606/Domain: fibronectin type III repeat homology <FN3A>
F:635-695/Domain: fibronectin type III repeat homology <FN3B>
F:742-739/Domain: transmembrane #status predicted <TM>
F:740-858/Domain: intracellular #status predicted <INT>
F:41-96;139-189;235-288;330-396;437-490/Disulfide bond: #status predicted
F:222.316,348,434,460,489/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 101.5; DB 1; Length 858;
Best Local Similarity 18.0%; Pred. No. 11;
Matches 43; Conservative 42; Mismatches 99; Indels 55; Gaps 8;

QY 3 RG-VPRHLLVLDLALLPATQGNKYVLGKKGPTVLLTCAQSKSIQFHW-KNSNQIK 60
DB 196 RGEINPKQIVYVNPPTVQARRQSIYNATANLGGSVTLVCDADGPEPTMSWTQDGEPIE 255
QY 61 ILGNQGSFLTGPSPKLNDRADRSRLWDQGNFPIIKNLKIEDPTICEVE----DQKE 116
DB 256 -----NHEEDEGHIFSDDSSELTINVDKNDAEVCAENKAGQDA 299
QY 117 EVQLLVFG---LTPNSDTHLLGQSLTFTLESPGSSPSVQCRS----- 157
DB 300 SIHKVFPAKPIITYENQTALEIEQVTLTCEASGDEIPSTIWTSTRNISSEKASWTR 359
QY 158 PRGNKIQG-----KTLVSQLELDQSGWTC-----TVLQNKVKYEFKIDYPR 202
DB 360 PEKETLIDGHVNVSHARVSSLTKLSIQYTDAGSYICTASNTIGQDSQSMYLEVQVAPK 418

RESULT 86
TIMSNL
neural cell adhesion molecule 1 precursor, long domain splice form - mouse
N:Alternate names: NCAM-180
N:Contains: neural cell adhesion molecule, short domain splice form (NCAM-140)
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 31-Dec-2000
C:Accession: A29673; S00844; S00384; A28281; A44290; S00353
R:Barthelemy, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaux, J.C.; Hirsch, M.R.; Pontet, E.MBO J. 6, 907-914, 1987
A:Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000
A:Reference number: A29673; MUID:87246524; PMID:3595563
A:Accession: A29673
A:Molecule type: mRNA
A:Residues: 1-548; 'T', 550-571, 'T', 573-574, 'D', 576-588, 'WQPS', 593, 'S', 595-599, 'P', 601, 'L
A:Cross-references: EMBL:Y00051; NID:G53342; PID:CA68263.1; PID:G53343
R:Santoni, M.J.; Barthelemy, D.; Barbais, J.A.; Hirsch, M.R.; Steinmetz, M.; Goridis, C.; 'Nucleic Acids Res. 15, 8621-8641, 1987
A:Title: Analysis of cDNA clones that code for the transmembrane forms of the mouse neu
A:Reference number: S00844; MUID:88067687; PMID:3684567
A:Accession: S00844
A:Molecule type: mRNA

A:Residues: 529-809, 1077-1115 <SAN>
A:Cross-references: EMBL:X06328; NID:G53322; PIDN:CAA29641.1; PID:G817984
R:Barbas, J.A.; Chalk, J.C.; Seimetz, M.; Gortidis, C.
EMBO J. 7, 625-632, 1988
A:Title: Differential splicing and alternative polyadenylation generates distinct NCAM
A:Reference number: S00382; MUID:88263628; PMID:3396534
A:Accession: S00384
A:Molecule type: DNA
A:Residues: 642-1115 <BAR>
A:Cross-references: EMBL:X07195
R:Barthels, D.; Vopper, G.; Wille, W.
Nucleic Acids Res. 16, 4217-4225, 1988
A:Title: NCAM-180, the large isoform of the neural cell adhesion molecule of the mouse.
A:Reference number: A28281; MUID:88247737; PMID:2454455
A:Accession: A28281
A:Molecule type: mRNA
A:Residues: 804-1081 <BA3>
A:Cross-references: EMBL:X07244; NID:G53321; PIDN:CAA30230.1; PID:G929720
R:Rougon, G.; Marshak, D.R.
J. Biol. Chem. 261, 3396-3401, 1986
A:Title: Structural and immunological characterization of the amino-terminal domain of
A:Reference number: A44290; MUID:86140120; PMID:3512556
A:Accession: A44290
A:Molecule type: protein
A:Residues: 20-36 <ROU>
C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C:Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IMM
C:Genetics:
A:Gene: NCAM
A:Map position: 9
A:Intons: 643/3; 701/1; 770/2; 809/2; 1076/2
C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-115/Domain: Product: neural cell adhesion molecule, long domain splice form #status experi
F:20-809,1077-1115/Product: neural cell adhesion molecule, short domain splice form #sta
F:20-711/Domain: extracellular #status predicted <EXT>
F:34-98/Domain: immunoglobulin homology <IMM1>
F:132-191/Domain: immunoglobulin homology <IMM2>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-290/Domain: immunoglobulin homology <IMM3>
F:262-272/Region: NCAM binding #status predicted
F:323-388/Domain: immunoglobulin homology <IMM4>
F:420-482/Domain: immunoglobulin homology <IMM5>
F:519-596/Domain: fibronectin type III repeat homology <FN3A>
F:625-685/Domain: fibronectin type III repeat homology <FN3B>
F:712-729/Domain: transmembrane #status predicted <TM>
F:770-1115/Domain: intracellular #status predicted <INT>
F:41-96,119-189,235-288,330-366,427-480/Diulfide bonds: #status predicted
F:222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.7%; Score 101.5; DB 1; Length 1115;
Best Local Similarity 19.9%; Pred. No. 15;
Matches 46; Conservative 42; Mismatches 94; Indels 49; Gaps 10;

3 RG-VPRHLLVQLALPLPAATGKQKVVLTCTASOKSIQFHW-KNSNQIK 60
196 RGEINFDIOIVIVNPPTVAPROSIVATNLTGQSVTLVADAGFPPTSMWKDGP 255
61 ILGNQGFLLTKGPKLNDRAD--SRSLMDQGNPPLIKLKIEDSDTYICEV---DQ 114
256 -----NEEDERSRSSVSSSE--VITRVNDKNDLDEYVICENKAKGEQ 237
115 KEEVOLLVFG---LTNSDTHLLQGGSLTLTESPPGSSPVQCR-----SPRKN 164
298 DASHLKVFAPKPKITTYENGTAMELEQVLTLCASDPIPSITWRSTNNTSSEQD 357
165 GG-----KTLVSQLELDGSGTTC---TVLONOKKVEKIDIVR 202
358 GHNVVSRHARVSSLTLKSIQYRDAGEYWCASNTIGDSIDLEFYARK 408

RESULT 87
S10004
hypothetical protein 6 - fowl adenovirus 1
C:Species: Aviatadenovirus gall (fowl adenovirus 1, CEL0)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Oct-1999
C:Accession: S10004
R:Akopian, T.A.; Kruglyak, V.A.; Rikina, M.B.; Naroditsky, B.S.; Tikhonenko, T.I.
Nucleic Acids Res. 18, 2825, 1990
A:Title: Sequence of an avian adenovirus (CEL0) DNA fragment (0-11.2%).
A:Reference number: S10004; MUID:90251474; PMID:2160072
A:Accession: S10004
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-135 <AKO>
A:Cross-references: EMBL:X17217; NID:G58537; PIDN:CAA35086.1; PID:G58538

Query Match 7.7%; Score 101; DB 2; Length 135;
Best Local Similarity 27.5%; Pred. No. 1.3;
Matches 38; Conservative 21; Mismatches 65; Indels 14; Gaps 4;

11 LVVLQALPLPAAT--QGNKVLGKGGDTVELTCTASOKSIQFHWKNSNQIKLNGSGF 68
1 MLTLTVLLVGYTLADHPTLVAPKGGSTELGVAKQKQYFWRFGN-LKIVIAEMSS 59
69 LTKGPKLNDRADSRSLMDQGNPPLIKLKIEDSDTYICEVQKEVQLVFLGTAN 128
60 TMLKTFPNDGNFQNRSENPPTKHLNLTINASVDSGTSLSHOENDG-----TEH 110

129 SDTH--LLQGOSLTLE 144
111 TDFKVIQVGMSTLYTLQ 128

RESULT 88
S22383
axinin 1 precursor - chicken
N:Alternate names: neural cell adhesion molecule AxCAM
C:Species: Gallus gallus (chicken)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: S22383; S34107; S69332; S22128
R:Zuellig, R.A.; Rader, C.; Schreeder, A.; Kalousek, M.B.; von Bohlen und Halbach, F.;
Eur. J. Biochem. 204, 453-463, 1992
A:Title: The axonally secreted cell adhesion molecule, axonin-1. Primary structure, imm
A:Reference number: S22383; MUID:92174898; PMID:1311675
A:Accession: S22383
A:Molecule type: mRNA
A:Residues: 1-1036 <ZUE1>
A:Cross-references: EMBL:X63101; NID:G62852; PIDN:CAA44815.1; PID:G62853
A:Accession: S34107
A:Molecule type: protein
A:Residues: 29-49,51-80;84-95;100-117;120-128;130-141;143-176;243-254;256-296;303-336;33
R:Giger, R.U.; Vogt, L.; Zuellig, R.A.; Rader, C.; Henahan-Beatty, A.; Wolter, D.P.; So
Eur. J. Biochem. 227, 617-628, 1995
A:Title: The gene of chicken axonin-1. Complete structure and analysis of the promoter.
A:Reference number: S69332; MUID:95172044; PMID:7867620
A:Accession: S69332
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1002-1036 <GIG>
A:Cross-references: EMBL:X79607
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
C:Superfamily: contactin; fibronectin type III repeat homology; immunoglobulin homology
C:Keywords: cell adhesion
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-1036/Product: axonin 1 #status predicted <MAT>
F:336-392/Domain: immunoglobulin homology <IMM>

Query Match 7.7%; Score 101; DB 2; Length 1036;
Best Local Similarity 23.2%; Pred. No. 15;
Matches 53; Conservative 30; Mismatches 87; Indels 58; Gaps 9;

34 GDTVELCTASOKKSIQFHWKNSNQIKLNGSGFLLTKGPKLNDRADSRSLMDQGNFP 93

F:23-88/Disulfide bonds: #status predicted

Query Match 7.6%; Score 100.5; DB 1; Length 108;

Best Local Similarity 39.0%; Pred. No. 1.1;

Matches 30; Conservative 5; Mismatches 37; Indels 5; Gaps 2;

QY 34 GDTVELTCTSOXKSIQIFHKNSNQIKILNGSFLTKGSKLNDPDRSRSLMDQG-NF 92

DB 16 GDTVTMTCSQSGSTINLNFQOKP---CKAPKLLYGASILEDGVPFRFSSRRYGTDF 71

QY 93 PLIIRKLIKEDSDTYIC 109

DB 72 TLTISSLEDDMATYFC 88

RESULT 91

C42632

cell adhesion molecule apCAM (clone d12) - California sea hare

C/Species: Aplysia californica (California sea hare)

C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000

C/Accession: C42632

R:Mayford, M.; Barzilai, A.; Keller, F.; Schacher, S.; Kandel, E.R.

Science 256, 638-644, 1992

A/Title: Modulation of an NCAM-related adhesion molecule with long-term synaptic plastic

A/Reference number: A42632; PMID:92263095; PMID:1585176

A/Accession: C42632

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: nucleic acid

A/Residues: 1-765 <MAY>

A/Experimental source: CNS

A/Note: sequence extracted from NCBI backbone (NCBIP:101351)

C/Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu

Query Match 7.6%; Score 100.5; DB 2; Length 765;

Best Local Similarity 21.4%; Pred. No. 11;

Matches 50; Conservative 34; Mismatches 79; Indels 71; Gaps 11;

QY 29 VLGKGDVELTCTASOKSIOFHKNSNQIKILNGSFLTKGSKLNDPDRSRSLMD 88

DB 136 ILGEGE--EVEGVSGKPAFTVTKFENNTKIEAGE---KYTALN----- 176

QY 89 QGNFPLIIRKLIKEDSDTYICEV-----EDQKEEVLVFGI-----TANSDTHLQ 135

DB 177 ----KLIIDLSIEDTKKYLCDIIVDTGETKDFIDFTVVKLPTALPTIHPDNPXV- 231

QY 136 GQSLTLTLESPPGSSPSVQCRS---PRKNIQGGKTLVSQLELQDSGTWCTV----- 186

DB 232 GDEVKITCOATGVPPPTQFKGDMVWVTDENVNG-VLTINPLKTTDQATYTCIATNKG 290

QY 187 -LQNGKVEFKIDIVPRASALPAPT-----GSALPDPOFASALPDP 227

DB 291 FAESNTLTVK-----PRTIEMERTYDAVSGELITTTAKGDPEP 333

RESULT 92

A35963

protein-tyrosine kinase (EC 2.7.1.112) cek2 precursor - chicken

C/Species: Gallus gallus (chicken)

C/Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 16-Jul-1999

C/Accession: A35963

R:Pasquale, E.B.

A/Title: A distinctive family of embryonic protein-tyrosine kinase receptors.

A/Reference number: A35963; PMID:90332672; PMID:2165604

A/Accession: A35963

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-806 <PAS>

A/Cross-references: GB:M35195; NID:g211442; PIDN:AAA48664.1; PID:g211443

C/Genetic: 82

C/Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote

C/Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-806/Product: protein-tyrosine kinase cek2 #status predicted <MAT>

F:24-368/Domain: extracellular #status predicted <EXT>

F:131-138/Region: acidic

F:262-335/Domain: immunoglobulin homology <IMM>

F:369-389/Domain: transmembrane #status predicted <TM>

F:390-806/Domain: intracellular #status predicted <INT>

F:464-749/Domain: protein kinase homology <KIN>

F:472-480/Region: protein kinase ATP-binding motif

F:51-107,170-222,269-333/Disulfide bonds: #status predicted

F:96,219,256,288,309,322/Binding site: carboxylate (Asn) (covalent) #status predicted

F:502,519,611/Active site: Lys, Glu, Asp #status predicted

F:516,629/Binding site: magnesium (Asn, Asp) #status predicted

F:642/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 7.6%; Score 100.5; DB 2; Length 806;

Best Local Similarity 20.8%; Pred. No. 12;

Matches 59; Conservative 32; Mismatches 85; Indels 107; Gaps 12;

QY 10 LLLVQLALPANTG-----NKVLLKGGTVELTCTASQKSI 49

DB 11 LCLAAVAGLPAARRRGAERSGGQAAEYRSETAPLEELVFG-SGDTIELSCN-TGSSSV 68

QY 50 QFHW-----XNSNQIKILNGSFLTKG 72

DB 69 SVFWFKDGIAPSNRTHIGQKLKILNVSVDGLYSCKPRHSN-VLGNFTVAVTDS 126

QY 73 PSKLNDRADSRSLMDQGNFPLIIRKLIKEDSDTYICEVEDQKEEVLVFGILTANSDTH 132

DB 127 PSSGDEDDDDDES-----EDTGVPEWTPRDPKMEKLLAV--PAAN---- 164

QY 133 LQGGSLTLTLESPPGSSPSVOC-----RSPRGKNIQGGKTLVSQLEL-----OPD 179

DB 165 -----TVFRCPAGGNFTPTITLWLNKKEKRGHRIGIKLRHQMSLWESVPSDR 217

QY 180 GTWCTVQLQNKVE--FKDIVPRASALPAPTGSALPDPOF 220

DB 218 GNVTCVENKYNIGNIRRTQGLVLSRPHRIQAG--LPANQT 258

RESULT 93

B42632

cell adhesion molecule apCAM (clone d15) - California sea hare

C/Species: Aplysia californica (California sea hare)

C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Mar-2000

C/Accession: B42632

R:Mayford, M.; Barzilai, A.; Keller, F.; Schacher, S.; Kandel, E.R.

Science 256, 638-644, 1992

A/Title: Modulation of an NCAM-related adhesion molecule with long-term synaptic plastic

A/Reference number: A42632; PMID:92263095; PMID:1585176

A/Accession: B42632

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: nucleic acid

A/Residues: 1-812 <MAY>

A/Experimental source: CNS

A/Note: sequence extracted from NCBI backbone (NCBIP:101346)

C/Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu

Query Match 7.6%; Score 100.5; DB 2; Length 812;

Best Local Similarity 21.4%; Pred. No. 12;

Matches 50; Conservative 34; Mismatches 79; Indels 71; Gaps 11;

QY 29 VLGKGDVELTCTASOKSIOFHKNSNQIKILNGSFLTKGSKLNDPDRSRSLMD 88

DB 136 ILGEGE--EVEGVSGKPAFTVTKFENNTKIEAGE---KYTALN----- 176

QY 89 QGNFPLIIRKLIKEDSDTYICEV-----EDQKEEVLVFGI-----TANSDTHLQ 135

DB 177 ----KLIIDLSIEDTKKYLCDIIVDTGETKDFIDFTVVKLPTALPTIHPDNPXV- 231

QY 136 GQSLTLTLESPPGSSPSVQCRS---PRKNIQGGKTLVSQLELQDSGTWCTV----- 186

DB 232 GDEVKITCOATGVPPPTQFKGDMVWVTDENVNG-VLTINPLKTTDQATYTCIATNKG 290

Db 6034 NNNCRLLPSSRTDQKVKVHLEAEFISKQKLEEGEKAFCVCSISKEFP-VQMKRD-D 6091

QY 161 KNIQCG-----KTLVSQLEIDSGTMTCTV 186
 Db 6092 KTLSEGDYDVADGKKRVLVVKDATTQDMGTVMV 6128

RESULT 96

C30127
 N:transmembrane carcinoembryonic antigen 3 precursor - human
 N:Alternate names: CD66 splice form BQPC
 C:Species: Homo sapiens (man)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 23-Jul-1999
 R:Accession: C30127, I52597
 R:Barrett, T.R.; Krebschmer, A.; Austen, D.A.; Goebel, S.J.; Hart, J.T.; Elting, J.J.; J. Cell Biol. 108, 267-276, 1989
 A:Title: Carcinoembryonic antigens: alternative splicing accounts for the multiple mRNA
 A:Reference number: A92752; MUID:89139550; PMID:2537311
 A:Accession: C30127
 A:Molecule type: mRNA

A:Residues: 1-464 <BAR>
 A:Cross-references: EMBL:X16356; EMBL:X14784
 R:Matl, S.M.; Fawcett, J.; Murdoch, S.J.; Teixeira, A.M.; Gschmeisner, S.E.; Hajjbagheri Blood 84, 200-210, 1994
 A:Title: CD66 identifies the biliary glycoprotein (BGP) adhesion molecule: cloning, exp

A:Reference number: I52597; MUID:94289702; PMID:8018919
 A:Accession: I52597
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-464 <RES>

A:Cross-references: GB:571326; NID:9550030; PIDN:AA31183.1; PID:9550031
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
 C:Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protein
 F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-464/Product: transmembrane carcinoembryonic antigen 3 #status predicted <MAT>
 F:160-217/Domain: immunoglobulin homology <IMM1>
 F:252-301/Domain: immunoglobulin homology <IMM2>
 F:351-398/Domain: immunoglobulin homology <IMM3>
 F:44-455/Domain: transmembrane #status predicted <TM>
 F:104,111,115,152,182,197,208,224,232,254,274,288,292,302,309,345,351,363,378,405/Bindin

Query Match 7.6%; Score 100; DB 2; Length 464;

Best Local Similarity 18.6%; Pred. No. 6.8;

Matches 49; Conservative 35; Mismatches 78; Indels 102; Gaps 10;

QY 32 KKGDTVELCTASQKSIQFMKNSNOIKLNGSFLTKPSKLNDRADSRSLMDQGN 91

Db 250 RFGANLTLSCVAASNPDAQSW-----LNG-----TFQOST 281

QY 92 FPLIKRLKIEDSTYICEVDO-----KEVQLVFGLTANS 129

Db 282 QEFPIPIYVNSSYCHANNSTVGNRTTKTIVTELSPVAKQIKASKTTVIGDK 341

QY 130 DTHLQGSLLTLSPSSPSVQCR-----SPRKNIOGKTLVSQLEIDS 179

Db 342 D-----SVNLHST---NDTGISIMFKNQSLPSSEMKKSQNTTISINPKEDA 391

QY 180 GTWCTVLQNKQKVEKIDIVPRASALPAPPTGALPDP---QTASALPDPAPASALPA 235

Db 392 GTWCEVFN-----PISKQSDPIMLVNVALPQENGLS--PG 428

QY 236 ALA-----VTSFLGLGLGAVACVL 254

Db 429 AIAIGIVGVALVALVALACFL 452

RESULT 97

A32164
 biliary glycoprotein 1 precursor, splice form a - human
 N:Alternate names: transmembrane carcinoembryonic antigen 1 (TM1-CEA); transmembrane can
 N:Contents: biliary glycoprotein 1, splice form b; biliary glycoprotein 1, splice form x
 C:Species: Homo sapiens (man)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: A32164; A30127; B30127; A48078; S45664; S65539; A30847; G44476
 R:Hindoa, Y.; Neumaier, M.; Hefta, S.A.; Drzeniek, Z.; Wagener, C.; Shively, L.; Hefta, Proc. Natl. Acad. Sci. U.S.A. 86, 1668, 1989
 A:Reference number: A32164
 A:Contents: extratum

A:Accession: A32164
 A:Molecule type: mRNA

A:Residues: 1-526 <HIN>
 A:Cross-references: GB:U03858; NID:9179439; PIDN:AAA51826.1; PID:9179440
 R:Hindoa, Y.; Neumaier, M.; Hefta, S.A.; Drzeniek, Z.; Wagener, C.; Shively, L.; Hefta, Proc. Natl. Acad. Sci. U.S.A. 85, 6959-6963, 1988

A:Title: Molecular cloning of a cDNA coding biliary glycoprotein I: Primary structure of
 A:Reference number: A94206; MUID:88320555; PMID:2457922

A:Contents: annotation
 A:Note: the sequence shown in this reference has been completely corrected in reference
 R:Barrett, T.R.; Krebschmer, A.; Austen, D.A.; Goebel, S.J.; Hart, J.T.; Elting, J.J.; J. Cell Biol. 108, 267-276, 1989

A:Title: Carcinoembryonic antigens: alternative splicing accounts for the multiple mRNA.
 A:Reference number: A92752; MUID:89139550; PMID:2537311
 A:Accession: A30127
 A:Molecule type: mRNA

A:Residues: 1-526 <BAR>
 A:Cross-references: EMBL:X16354; NID:937197; PIDN:CAA34404.1; PID:937198; EMBL:X14784
 A:Experimental source: splice form a

A:Accession: B30127
 A:Molecule type: mRNA
 A:Residues: 1-319; 'D', 417-526 <BAR2>

A:Cross-references: EMBL:X14831; NID:937199; PIDN:CAA32940.1; PID:937200; EMBL:X14784
 A:Experimental source: splice form b
 R:Barrett, T.R.; Drake, L.; Pickle II, W. Mol. Cell. Biol. 13, 1273-1282, 1993

A:Title: Human biliary glycoprotein gene: characterization of a family of novel alterna
 A:Reference number: A48078; MUID:93140765; PMID:8423792
 A:Accession: A48078
 A:Molecule type: mRNA

A:Residues: 124-141; 'H', 417-526 <BAR3>
 A:Cross-references: GB:M76742; NID:9179480; PIDN:AAA57142.1; PID:9179481
 A:Experimental source: splice form x

A:Note: sequence extracted from NCBI backbone (NCBIN:123602, NCBI:123606)
 A:Note: neither the complete nucleic acid sequence nor the complete translation are sho
 R:Huck, W.; Neddelic, P.; Turbide, C.; Stanners, C.P.; Barnett, T.R.; Beauchemin, N. Eur. J. Biochem. 223, 529-541, 1994

A:Title: Transcriptional control of the human biliary glycoprotein gene, a CEA gene fam
 A:Reference number: S45664; MUID:94333343; PMID:8055523
 A:Accession: S45664
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-21 <HAU>

A:Cross-references: EMBL:X67277; NID:929447; PIDN:CAA47694.1; PID:9606777
 R:Neddelic, P.; Turbide, C.; Beauchemin, N. Eur. J. Biochem. 231, 104-114, 1995
 A:Title: Characterization and transcriptional activity of the mouse biliary glycoprotei

A:Reference number: S65939; MUID:95354678; PMID:7628460
 A:Accession: S65939
 A:Status: preliminary; translation not shown

A:Molecule type: DNA
 A:Residues: 1-21 <NED>

A:Cross-references: EMBL:X67277; NID:929447; PIDN:CAA47694.1; PID:9606777
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
 A:Note: only a part of the coding sequence is given

R:Khan, W.N.; Fraenkel, L.; Teglund, S.; Israelsson, A.; Bremer, K.; Hammarstrom, S. Genomics 14, 384-390, 1992
 A:Title: Identification of three new genes and estimation of the size of the carcinoemb
 A:Reference number: A44476; MUID:93052339; PMID:1427854
 A:Contents: annotation; alignment of related sequences

A:Gene: GDB:BG
 A:Gene: GDB:BG

A:Cross-references: GDB:127992; OMIM:109770
 A:Map position: 19q13.2-19q13.2
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
 C:Keywords: alternative splicing; glycoprotein; surface antigen; transmembrane protein
 F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>

F:1-34/Domains: signal sequence #status predicted <SIG>
 F:35-526/Product: biliary glycoprotein 1, splice from a #status predicted <MAY>
 F:35-428/Domains: extracellular #status predicted <EXT>
 F:35-319,'D',417-526/Product: biliary glycoprotein 1, splice form b #status predicted <V>
 F:35-141,'H',417-526/Product: biliary glycoprotein 1, splice form x #status predicted <V>
 F:160-217/Domains: immunoglobulin homology <IMM1>
 F:252-301/Domains: immunoglobulin homology <IMM2>
 F:341-398/Domains: immunoglobulin homology <IMM3>
 F:425-454/Domains: intracellular #status predicted <INT>
 F:453-526/Domains: intracellular #status predicted <INT>
 F:104,111,115,152,182,197,208,224,232,254,274,288,292,302,309,345,351,363,378,405,475/B1

Query Match 7.6%; Score 100; DB 1; Length 526;
 Best Local Similarity 18.6%; Pred. No. 7.8;

Matches 49; Conservative 35; Mismatches 78; Indels 102; Gaps 10;

QY 32 KKGDVLTCTAASOKSIOFHKNSNOIKLGNOSFLTKGPSKLNDRADSRSLMDQGN 91
 DB 250 RFGANLSTSCYASNPAAQYGM-----LING-----TFQOST 281
 QY 92 FPLIIKLIKIDSDTYICEVEDQ-----KEVQLLVFGLTANS 129
 DB 282 QELFIPNITVNNSSSYTCNRTTAKTIIITELSPVYAKPDIKASKTTVTGDK 341
 QY 130 DTHLLQGSLTLTLESPGSSPSVOCR-----SPRGNIOGKTLISVSOLELQDS 179
 DB 342 D-----SVNLTCST---NDTGISIRWFPKQSLPSEBRKLSQGNLTLSINRYKEDA 391
 QY 180 GTWTCVTLQNOQKVEFKIDIVRASALPAPPTGSALPDP---QTASALPDPAPASALPA 235
 DB 392 GTWCEVEN-----PISKQSDPIMLVNPNYVALPENGLS--PG 428
 QY 236 ALA-----VISFLGLGIGVACVL 254
 DB 429 AIAIGIVGVALVALIAVALACFL 452

RESULT 98

S33161
 IG kappa chain - sheep
 C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
 C/Accession: S33161
 R/Foley, R.C.; Beh, K.J.
 submitted to the EMBL Data Library, July 1990
 A/Description: Isolation and characterization of sheep kappa light chain cDNA.
 A/Reference number: S33161
 A/Accession: S33161
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-230 <FOL>
 A/Cross-references: EMBL:X54110; NID:G297103; PIDN:CAA38046.1; PID:G13642221
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotrimer; immunoglobulin
 F:143-212/Domains: immunoglobulin homology <IMV>

Query Match 7.5%; Score 99; DB 2; Length 230;
 Best Local Similarity 24.5%; Pred. No. 3.5;

Matches 53; Conservative 24; Mismatches 95; Indels 44; Gaps 8;

QY 10 LVLVQLALPRA-----TGKNVYLGKKDVTLELTCTASOKSIOFHKNSNOIKILG 63
 DB 2 LIGLILMLPGACDIOVTQSPSSLSLTERVSTICRTSQSVSNYLMYQOKP-----G 57
 QY 64 NQGSFLTKGPSKLNDRADSRSLMDQGNFPLIIKLIKIDSDTYIC-----109
 DB 58 QAPKLIIYATRLHTDVSRSGSGSDYTLTISNEADNTATYYCYQVESTPLARGG 117
 QY 110 -EVDQKEVQLLVFGLTANS DTHLLQGSLTLTLES---PPGSSPSVOCRSPRGNIO 164
 DB 118 TNVEIKSDAQPSVF-LFKPSEBQLRTGTVSVCCLVNDVFPKIDINVKYKDGVTQNSNFQ 176
 QY 165 G-----GKTLISVQLELQDSGTWTCV 186

DB 177 NSFNDQDSKSTYSLSSTLTLSSSEYQSHNAVCEV 212

RESULT 99

S40317
 IG kappa chain - human
 C/Species: Homo sapiens (man)
 C/Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
 C/Accession: S40317
 R/Klein, R.; Jaenichen, R.; Zachau, H.G.
 Eur. J. Immunol. 23, 3248-3271, 1993
 A/Title: Expressed human immunoglobulin chi genes and their hypermutation.
 A/Reference number: S40312; MUID:94080891; PMID:8258341
 A/Accession: S40317
 A/Status: preliminary; translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-129 <KLE>
 A/Cross-references: EMBL:X72427; NID:G441322; PIDN:CAA51095.1; PID:G441323
 C/Superfamily: immunoglobulin V region; immunoglobulin homology
 C/Keywords: heterotrimer; immunoglobulin
 F:37-111/Domains: immunoglobulin homology <IMV>

Query Match 7.5%; Score 98.5; DB 2; Length 129;
 Best Local Similarity 33.6%; Pred. No. 1.9;

Matches 36; Conservative 7; Mismatches 53; Indels 11; Gaps 3;

QY 11 LVLVQLALPRA-----TGKNVYLGKKDVTLELTCTASOKSIOFHKNSNOIKILGN 64
 DB 8 LIGLILMLPGACDIOVTQSPSSLSLTERVSTICRTSQSVSNYLMYQOKP-----GK 63
 QY 65 QGSFLTKGPSKLNDRADSRSLMDQGNFPLIIKLIKIDSDTYICE 110
 DB 64 APKFLIYAGASSLQSGVPSRSGSGSDYTLTISLQPEDPATYCCQ 110

RESULT 100

A34815
 carcinoembryonic antigen family member CCM6 precursor - human
 N/Alternate names: carcinoembryonic antigen family member NCA-95; CD67; nonspecific cro.
 C/Species: Homo sapiens (man)
 C/Date: 13-Jul-1990 #sequence_revision 02-Aug-1996 #text_change 20-Jun-2000
 C/Accession: S13524; I44476; A34815
 R/Berling, B.; Kolbinger, F.; Grunert, F.; Thompson, J.A.; Brombacher, F.; Buchegger, F.
 Cancer Res. 50, 6534-6539, 1990
 A/Title: Cloning of a carcinoembryonic antigen gene family member expressed in leukocytes.
 A/Reference number: S13524; MUID:91003998; PMID:2208113
 A/Accession: S13524
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-349 <BER>
 A/Cross-references: EMBL:X52378; NID:G29918; PIDN:CAA36604.1; PID:G29919
 R/Xian, W.N.; Frengemyr, L.; Teglund, S.; Israelsson, A.; Bremer, K.; Hammarstrom, S.
 Genomics 14, 384-390, 1992
 A/Title: Identification of three new genes and estimation of the size of the carcinoemb.
 A/Reference number: A44476; MUID:93052339; PMID:1427854
 A/Accession: I44476
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 35-141 <KHA>
 R/Arakawa, F.; Kuroki, M.; Mitsu, Y.; Oikawa, S.; Nakazato, H.; Matsuo, Y.
 Biochem. Biophys. Res. Commun. 166, 1063-1071, 1990
 A/Title: Characterization of a cDNA clone encoding a new species of the nonspecific cro.
 A/Reference number: A34815; MUID:90165902; PMID:2306228
 A/Accession: A34815
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-113,'R',115-321,'V',323-349 <ARA>
 A/Cross-references: GB:D90064; NID:G219935; PIDN:BA41108.1; PID:G219936
 A/Experimental source: white blood cells
 C/Genetics:
 A/Gene: GDB:CCM6
 A/Cross-references: GDB:127667

A:Map position: 19q13.2-19q13.2
 C:Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-terminal
 C:Keywords: phosphatidylinositol 11kage
 F:1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>
 F:160-217/Domain: immunoglobulin homology <IMM1>
 F:252-301/Domain: immunoglobulin homology <IMM2>

Query Match 7.5%; Score 98.5; DB 2; Length 349;

Best Local Similarity 20.2%; Pred. No. 6.2;

Matches 60; Conservative 51; Mismatches 87; Indels 99; Gaps 15;

```

Qy 5 VPRHLLLVQL-----ALPAAQGNKVL-----GKKGDT 36
   : : : : : : : : : : : : : : : : : : : : : :
Db 13 IPWQGLLITSLTFMNPPTTAQLITEAVPSNAEGKEVLLVHNLPQDPKGYWYKGET 72
   : : : : : : : : : : : : : : : : : : : : : :
Qy 37 VELCTASQKKSIOFHWKNSNQIKILGNQGSFLTKGPSKLNDRADSRRLMDQGNFPLII 96
   : : : : : : : : : : : : : : : : : : : : : :
Db 73 VD-----ANRRIIGY-----VISNQ--ITPGPAYSN---RETIYP--NASLMM 109
   : : : : : : : : : : : : : : : : : : : : : :
Qy 97 KNLKIEDSDTY---ICEVEDQKEV--QLLVF-----GITANSDFHLQGSLLTITLES 145
   : : : : : : : : : : : : : : : : : : : : : :
Db 110 RNVTKNDTGSYTLQVTKLNLMSSEVVGQFSVHPETPKPSISSNNSNPFVEDKDAVAFTCEP 169
   : : : : : : : : : : : : : : : : : : : : : :
Qy 146 PP-----GSSPSVQCRSPRGKNIQSGKTLVSQLELDPSGTWCTVLONOKKVER 195
   : : : : : : : : : : : : : : : : : : : : : :
Db 170 ETQNTTYLWMVNGOSLPV---SPRLQLSNGNRITLILSVTRANDVGYECEI-QN----- 219
   : : : : : : : : : : : : : : : : : : : : : :
Qy 196 KIDIVPRASALPAPPTGSALPDQOTASALPDPPAASALPALAVISFLGLGLGVAC 252
   : : : : : : : : : : : : : : : : : : : : : :
Db 220 -----PASANFSDPVTNLVLYGPDAPFISPSD---TYHAGVNLNLISC 259
   : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: August 3, 2004, 13:15:07
 Job time : 9.92937 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 3, 2004, 12:50:44 ; Search time 3.5436 Seconds

(without alignments)
3791.557 Million cell updates/sec

Title: SEQ6
Perfect score: 1317
Sequence: 1 MNRGVPRHLLVLQLALP.....VISFLGLGVACVLARTR 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 125 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1023	77.7	458	1 CD4_HUMAN	P01730 homo sapien
2	999	75.9	458	1 CD4_PANTR	P16004 pan troglod
3	912	69.2	458	1 CD4_MACFA	P79188 macaca fasc
4	910	69.1	458	1 CD4_MACRU	P79184 macaca fusc
5	904	68.6	458	1 CD4_MACMU	P16003 macaca mula
6	903	68.6	458	1 CD4_MACNE	Q08340 macaca neme
7	885	67.2	458	1 CD4_CERAE	Q08336 cercopithec
8	784	59.5	397	1 CD4_CERYA	Q08339 erythrocebu
9	783	59.5	397	1 CD4_BRYPA	Q29007 saimiri bci
10	737.5	56.0	457	1 CD4_SAIISC	P33705 canis famli
11	598	45.4	463	1 CD4_CANPA	P46630 oryctolagus
12	596.5	45.3	459	1 CD4_RABIT	P05540 rattus norv
13	490.5	37.2	457	1 CD4_RAT	P05332 mus musculu
14	475	36.1	457	1 CD4_MOUSE	P09564 homo sapien
15	332	25.2	240	1 CD7_HUMAN	O05695 rattus norv
16	133	10.1	1259	1 CAME_RAT	P11627 mus musculu
17	131	9.9	1260	1 CAME_MOUSE	Q31449 homo sapien
18	129	9.8	338	1 LAMP_HUMAN	O35136 mus musculu
19	127	9.6	837	1 NCN2_MOUSE	O62813 rattus norv
20	126	9.6	338	1 LAMP_RAT	P29534 rattus norv
21	124	9.4	739	1 VCAL_RAT	Q09206 rattus norv
22	122	9.3	348	1 KILO_RAT	Q26474 echinocerc
23	120	9.1	349	1 LACH_SCHAM	Q08902 figu rubrip
24	120	9.1	1277	1 CAME_FUGRU	P14781 gallus gall
25	118	9.0	1010	1 CONT_CHICK	Q04781 brachydanio
26	118	9.0	1197	1 CAME_BRARE	P11834 bos taurus
27	117.5	8.9	345	1 OPCM_BOVIN	Q088919 gallus gall
28	117	8.9	338	1 LAMP_CHICK	P19320 homo sapien
29	117	8.9	739	1 VCAL_HUMAN	Q41982 homo sapien
30	116.5	8.8	345	1 OPCM_HUMAN	Q24372 drosophila
31	116.5	8.8	359	1 LACH_DROME	P01832 cyccolagus
32	112.5	8.5	773	1 PIGR_RABIT	P06731 homo sapien
33	111	8.4	702	1 CEAS_HUMAN	

34	110.5	8.4	2029	1 LAR_DROME	P16621 drosophila
35	110	8.3	739	1 VCAL_MOUSE	P29533 mus musculu
36	109.5	8.3	333	1 AMAL_DROME	P15364 drosophila
37	109	8.3	1257	1 CAME_HUMAN	P32004 homo sapien
38	109	8.3	1367	1 VGR2_MOUSE	P35918 mus musculu
39	109	8.3	4391	1 PGBM_HUMAN	P98160 homo sapien
40	108	8.2	257	1 FCEA_HUMAN	P12319 homo sapien
41	108	8.2	319	1 A33_HUMAN	Q09795 homo sapien
42	108	8.2	837	1 NCN2_HUMAN	O15394 homo sapien
43	107	8.1	761	1 NCAL_HUMAN	P13592 homo sapien
44	107	8.1	848	1 NCAL_HUMAN	P13591 homo sapien
45	107	8.1	1447	1 DCC_MOUSE	P70211 mus musculu
46	107	8.1	3707	1 PGBM_MOUSE	O05793 mus musculu
47	106.5	8.1	345	1 OPCM_RAT	P32736 rattus norv
48	105	8.0	2012	1 DSCA_HUMAN	O60469 homo sapien
49	105	8.0	344	1 NTRI_HUMAN	Q29121 homo sapien
50	105	8.0	344	1 NTRI_MOUSE	Q29120 mus musculu
51	104	7.9	344	1 CONT_MOUSE	Q62718 rattus norv
52	104	7.9	1018	1 CONT_HUMAN	Q12860 homo sapien
53	104	7.9	1020	1 CONT_MOUSE	P12960 mus musculu
54	104	7.9	1021	1 CONT_RAT	O63198 rattus norv
55	103.5	7.9	108	1 KVS5_MOUSE	P01652 mus musculu
56	103.5	7.9	129	1 KVIW_HUMAN	P04431 homo sapien
57	103.5	7.9	210	1 CD7_MOUSE	P50283 mus musculu
58	103.5	7.9	1302	1 NRG_DROME	P20241 drosophila
59	103	7.8	519	1 ECTO_RAT	P16573 rattus norv
60	103	7.8	521	1 LAG3_MOUSE	O61790 mus musculu
61	102.5	7.8	862	1 CD22_MOUSE	P35329 mus musculu
62	102.5	7.8	1348	1 VGR2_COTJA	P52583 coturnix co
63	102	7.7	521	1 CEAL_MOUSE	P31809 mus musculu
64	102	7.7	1343	1 VGR2_RAT	O08775 rattus norv
65	101.5	7.7	509	1 SHS1_RAT	P07710 r protein-t
66	101.5	7.7	725	1 NCN2_MOUSE	P13594 mus musculu
67	101.5	7.7	858	1 NCAL_RAT	P13596 rattus norv
68	101.5	7.7	1115	1 NCAL_MOUSE	P13595 mus musculu
69	101	7.7	135	1 YOR6_ADEG1	P20748 avian adeno
70	101	7.7	1036	1 AXOI_CHICK	P28685 gallus gall
71	101	7.7	3375	1 UN52_CAMEL	O06551 caenorhabdi
72	100.5	7.6	108	1 KVSQ_MOUSE	P01650 mus musculu
73	100.5	7.6	337	1 OPCM_CHICK	Q08892 gallus gall
74	100.5	7.6	806	1 CEK2_CHICK	P18460 gallus gall
75	100	7.5	526	1 CEAL_HUMAN	P13668 homo sapien
76	98.5	7.5	349	1 CEAL_HUMAN	P11997 homo sapien
77	98.5	7.5	686	1 SILB_HUMAN	Q09167 homo sapien
78	98	7.4	209	1 CD8B_SAIISC	Q28877 saimiri bci
79	98	7.4	764	1 PIGR_HUMAN	P01833 homo sapien
80	98	7.4	880	1 TYO3_MOUSE	P55144 mus musculu
81	97.5	7.4	108	1 KVS1_MOUSE	P01653 mus musculu
82	97.5	7.4	332	1 ICOL_MOUSE	O91458 mus musculu
83	97.5	7.4	620	1 SMP_COTJA	O92154 coturnix co
84	97.5	7.4	771	1 PIGR_MOUSE	O70570 mus musculu
85	97.5	7.4	1333	1 VGR1_MOUSE	P35959 mus musculu
86	97.5	7.4	6632	1 UN89_CAMEL	O01761 caenorhabdi
87	97	7.4	822	1 FGRI_MOUSE	P16092 mus musculu
88	97	7.4	1091	1 NCAL_CHICK	P13590 gallus gall
89	97	7.4	1447	1 DCC_HUMAN	P43316 homo sapien
90	96.5	7.3	117	1 KVIU_HUMAN	P01602 homo sapien
91	96.5	7.3	769	1 PIGR_RAT	P15050 rattus norv
92	96.5	7.3	1694	1 SN_MOUSE	O62220 mus musculu
93	96	7.3	628	1 LU_HUMAN	P50895 homo sapien
94	95.5	7.3	1363	1 VGR3_MOUSE	P35917 mus musculu
95	95.5	7.3	1493	1 NEO1_MOUSE	P97798 mus musculu
96	95	7.2	245	1 FCEB_RAT	P12371 rattus norv
97	95	7.2	1356	1 VGR2_HUMAN	P35968 homo sapien
98	94.5	7.2	108	1 KVIIP_HUMAN	P01608 homo sapien
99	94.5	7.2	210	1 CD8B_PONPY	P30444 pongo pygma
100	94.5	7.2	524	1 BUTY_MOUSE	O62556 mus musculu
101	94	7.1	108	1 KVI1_HUMAN	P01611 homo sapien
102	94	7.1	808	1 FGRI_MOUSE	O03142 mus musculu
103	94	7.1	822	1 FGRI_HUMAN	P11362 homo sapien
104	94	7.1	1336	1 VGR1_RAT	P53767 rattus norv
105	93.5	7.1	824	1 MLTI_HUMAN	Q09498 homo sapien
106	93.5	7.1	1906	1 KMLS_CHICK	P11799 gallus gall

ID	CD4_HUMAN	STANDARD	PRT	458 AA	ALIGNMENTS
107	93	7.1	302	1	Y016_RICPR
108	93	7.1	393	1	IL15_CERAE
109	93	7.1	853	1	NCAL_BOVIN
110	93	7.1	1338	1	VGR1_HUMAN
111	92.5	7.0	108	1	KVSR_MOUSE
112	92.5	7.0	213	1	CD8B_MOUSE
113	92.5	7.0	344	1	CEA6_HUMAN
114	92.5	7.0	997	1	SCP1_RAT
115	92.5	7.0	1443	1	NEO1_CHICK
116	92	7.0	697	1	SILA_HUMAN
117	91.5	6.9	315	1	HEMA_VACCT
118	91.5	6.9	468	1	IL6A_HUMAN
119	91.5	6.9	738	1	PECA_HUMAN
120	91	6.9	639	1	NRG1_HUMAN
121	90.5	6.9	108	1	KVSM_MOUSE
122	90.5	6.9	335	1	PSG5_HUMAN
123	90.5	6.9	1567	1	FMN2_MOUSE
124	90.5	6.9	1814	1	TSC2_MOUSE
125	90	6.8	265	1	CEA7_HUMAN

ALIGNMENTS

RESULT 1

CD4_HUMAN

AC P01730; STANDARD; PRT; 458 AA.

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).

GN CD4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=85254948; PubMed=2990730;

RA Maddon P.J., Littman D.R., Godfrey M., Maddon D.E., Chess L., Axel R.;

RT "The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family.";

RL Cell 42:93-104(1985).

RN [2]

RP REVISION TO 26.

RX MEDLINE=89028665; PubMed=3263213;

RA Littman D.R., Maddon P.J., Axel R.;

RT "Corrected CD4 sequence.";

RL Cell 55:541-541(1988).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=96303695; PubMed=8723724;

RA Anasari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilly C.E., Spanos S., Malley T., Gibbs R.A.;

RT "A gene-rich cluster between the CD4 and triosephosphate isomerase genes at human chromosome 12p13.";

RL Genome Res. 6:314-326(1996).

RN [4]

RP SEQUENCE FROM N.A. AND VARIANT TRP-265.

RX MEDLINE=91216786; PubMed=1708753;

RA Hodge T.W., Sasso D.R., McDougal J.S.;

RT "Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240.";

RL Hum. Immunol. 30:99-104(1991).

RN [5]

RP SEQUENCE FROM N.A.

RX MEDLINE=22388257; PubMed=12477932;

RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D., Alteshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mollath S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Rodriguez Y.S.N., Krzywicki M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E., Jones S.J.W., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [6]

RP SEQUENCE OF 28-424 FROM N.A.

RC TISSUE=Blood;

RX MEDLINE=93049640; PubMed=1425921;

RA Fomsgaard A., Hirsch V.M., Johnson P.R.;

RT "Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus.";

RL Eur. J. Immunol. 22:2973-2981(1992).

RN [7]

RP SEQUENCE OF 26-394.

RX MEDLINE=90078232; PubMed=2592374;

RA Carr S.A., Henling M.E., Folea-Wasserman G., Sweet R.W., Anumula K., Barr J.R., Huddleston M.J., Taylor P.;

RT "Protein and carbohydrate structural analysis of a recombinant soluble CD4 receptor by mass spectrometry.";

RL J. Biol. Chem. 264:21286-21295(1989).

RN [8]

RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-208.

RX MEDLINE=91061881; PubMed=1701030;

RA Wang J., Yan Y., Garrett T.P., Liu J., Rodgers D.W., Garlick R.L., Tarr G.E., Husein Y., Reinherz E.L., Harrison S.C.;

RT "Atomic structure of a fragment of human CD4 containing two immunoglobulin-like domains.";

RL Nature 348:411-418(1990).

RN [9]

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-208.

RX MEDLINE=91061882; PubMed=2247146;

RA Ryu S.-E., Kwong P.D., Truneh A., Porter T.G., Arthos J., Rosenberg M., Dai X., Xuong N.-H., Axel R., Sweet R.;

RT "Crystal structure of an HIV-binding recombinant fragment of human CD4.";

RL Nature 348:419-426(1990).

RN [10]

RP X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF 26-388.

RX MEDLINE=97311402; PubMed=9168119;

RA Wu H., Kwong P.D., Hendrickson W.A.;

RT "Dimeric association and segmental variability in the structure of human CD4.";

RL Nature 387:527-530(1997).

RN [11]

RP PALMITOYLATION.

RX MEDLINE=92317088; PubMed=1618861;

RA Chise B., Rose J.K.;

RT "Identification of palmitoylation sites on CD4, the human immunodeficiency virus receptor.";

RL J. Biol. Chem. 267:13593-13597(1992).

CC -I- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.

CC -I- SUBUNIT: Associates with p56-lck.

CC -I- SUBCELLULAR LOCATION: Type I membrane protein.

CC -I- SIMILARITY: Contains 3 immunoglobulin-like C2-type domain.

CC -I- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.


```

CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD4 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd4.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12807; AAA3572.1; -
DR EMBL; U47924; AAB51309.1; -
DR EMBL; M35160; AAA16069.1; -
DR EMBL; BC025782; AAB25782.1; -
DR PIR; A90872; RWHUT4.
DR PDB; 1CDH; 30-APR-94.
DR PDB; 1CDI; 30-APR-94.
DR PDB; 3CD4; 31-OCT-93.
DR PDB; 1CDU; 01-APR-97.
DR PDB; 1CDU; 01-APR-97.
DR PDB; 1CDY; 01-APR-97.
DR PDB; 1WBR; 12-MAR-97.
DR PDB; 1WIO; 07-JUL-97.
DR PDB; 1WIP; 07-JUL-97.
DR PDB; 1WIO; 07-JUL-97.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1G9N; 27-DEC-00.
DR PDB; 1GCI; 19-SEP-01.
DR PDB; 1JL4; 19-SEP-01.
DR GlycoSuiteDB; P01730; -
DR GeneW; HGNC:1678; CD4.
DR MIM; 186940; -
DR GO; GO:0042101; C:T-cell receptor complex; NAS.
DR GO; GO:0015026; F:coreceptor activity; NAS.
DR GO; GO:0015029; F:internalization receptor activity; TAS.
DR GO; GO:0042289; F:MHC class II protein binding; NAS.
DR GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR GO; GO:0006955; P:immune response; NAS.
DR GO; GO:0009405; P:pathogenesis; TAS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; NAS.
DR GO; GO:0030217; P:T-cell differentiation; NAS.
DR GO; GO:0045058; P:T-cell selection; NAS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; NAS.
DR InterPro; IPR000973; CD4 TCRg.
DR InterPro; IPR007110; Ig_Like.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PRO0692; CD4TCRANTIGEN.
DR SMART; SMO0406; IGV_1.
DR PROSITE; PS50835; IG_Like; 1.
KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KM Immune response; Repeat; Signal; Lipoprotein; Palmitate; 3D-structure;
KM Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSLEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT VARIANT 265 265
R -> W (in OKT4-negative populations).

```

```

FT STRAND 27 32 /FTId=VAR_003906.
FT TURN 33 34
FT STRAND 37 39
FT TURN 44 45
FT STRAND 51 55
FT TURN 56 57

Query Match 77.7%; Score 1023; DB 1; Length 458;
Best Local Similarity 99.5%; Pred. No. 1,1e-70;
Matches 199; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALLPAATQGNKVVGGKDYELCTASQSKSIQFMHNSNQIK 60
DQ 1 MNRGVPRHLLVLTQALLPAATQGNKVVGGKDYELCTASQSKSIQFMHNSNQIK 60
QY 61 ILNGQGSFLTKGSPKLNDRADSRSLWDQGNFPLIKNKIEDSDTYICEVEDQKEVOL 120
DQ 61 ILNGQGSFLTKGSPKLNDRADSRSLWDQGNFPLIKNKIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTNSDTHLLOGLSLTLTSPSSPSPVQCRSPRKNIQGGKTLISVSOLELDSG 180
DQ 121 LVFGLTNSDTHLLOGLSLTLTSPSSPSPVQCRSPRKNIQGGKTLISVSOLELDSG 180
QY 181 TWICTVLOKOKVEFKIDIV 200
DQ 181 TWICTVLOKOKVEFKIDIV 200

RESULT 2
CD4_PANTR
ID CD4_PANTR STANDARD; PRT; 458 AA.
AC P16004;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN CD4.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9018264; PubMed=2107024;
RA Cameroni D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE OF 26-424 FROM N.A.
RX TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```
CC -----
DR EMBL; M31135; AAA35407.1; -.
DR EMBL; X73723; CAAS1749.1; -.
DR PIR; B32722; RWCZT4.
DR HSSP; P01730; IMIQ.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR000973; CD4 TCAG.
DR InterPro; IPR007110; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW SIGNAL; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 42 42
FT CONFLICT 62 62
FT CONFLICT 191 191
SQ SEQUENCE 458 AA; 51057 MW; AVCA3A8A527D3AD CRC64;

Query Match 75.9%; Score 999; DB 1; Length 458;
Best Local Similarity 97.5%; Pred. No. 7.6e-69;
Matches 195; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MNRGVPRHLILVQLALPAATQGNKVLGKGGDTVELCTASOKKSIOFHKNSNQIK 60
DB 1 MNRGVPRHLILVQLALPAATQGNKVLGKGGDTVELCTASOKKSIOFHKNSNQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVRELTRANSPTHTLQGSGLTLESPPGSSPSVOCSPRGNTQSGKTLVSQLELDQSG 180
DB 121 LVRELTRANSPTHTLQGSGLTLESPPGSSPSVOCSPRGNTQSGKTLVSQLELDQSG 180
QY 181 TWICTVLONOKKVEFKIDIV 200
DB 181 TWICTVLONOKKVEFKIDIV 200
QY 181 TWICTVLONOKKVEFKIDIV 200
DB 181 TWICTVLONOKKVEFKIDIV 200

RESULT 3
CD4_MACFA ID_CD4_MACFA STANDARD; PRT; 458 AA.
AC P79185;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).
```

```
GN CD4.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxId=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymocytes;
RA Tatemui M., Yabe M., Yamada Y.K.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63349; BAA09673.1; -.
DR HSSP; P01730; IMBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR000973; CD4 TCAG.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW SIGNAL; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
SQ SEQUENCE 458 AA; 50872 MW; 9105479F85C56FF7 CRC64;

Query Match 69.2%; Score 912; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 3.2e-62;
Matches 176; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

QY 1 MNRGVPRHLILVQLALPAATQGNKVLGKGGDTVELCTASOKKSIOFHKNSNQIK 60
DB 1 MNRGVPRHLILVQLALPAATQGNKVLGKGGDTVELCTASOKKSIOFHKNSNQIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120
```

```

Db 61 ILGIQGSFLTKGPSKLSDRADSRKSLMDQGCFSMTIKNLKIEDSDTYICEVKNKEEVEL 120
Qy 121 LVFGLTANSPTHLQGSGLTTLTSPSSPSVQCRSPRGKNIQCGKTLVSQLELDQSG 180
Db 121 LVFGLTANSPTHLQGSGLTTLTSPSSPSVQCRSPRGKNIQCGKTLVSQLELDQSG 180
Qy 181 TWCTCTVLONOKKVEFKIDIV 200
Db 181 TWCTCTVSODKTVKFKIDIV 200

RESULT 4
CD4_MACFU STANDARD; PRT: 458 AA.
AC P79184;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
CD4/Leu-3).
OS Macaca fuscata fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
NCBI_TaxID=9543;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (by similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63348; BAA09672.1; -.
DR HSSP; P01730; 1MBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:Immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007159; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; IG_V.
DR Pfam; PF00047; IG_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458 T-CELL SURFACE GLYCOPROTEIN CD4.
FT DOMAIN 26 396 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 397 418 POTENTIAL.
FT DOMAIN 419 458 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 125 IG-LIKE V-TYPE.
FT FT 126 203 IG-LIKE C2-TYPE 1.
FT DOMAIN 204 317 IG-LIKE C2-TYPE 2.

```

```

FT DOMAIN 318 374 IG-LIKE C2-TYPE 3. (POTENTIAL).
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 328 370 BY SIMILARITY.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 458 AA; 50828 MW; 76B3E7E0818535 CRC64;

Query Match 69.1%; Score 910; DB 1; Length 458;
Best Local Similarity 87.5%; Pred. No. 4.6e-62;
Matches 175; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MNRGVPFRHLILVQLALPPATQGNKYVLAKKDDYELTCTASQKKSIOFHKNKNOIK 60
Db 1 MNRGIPFRHLILVQLALPPAVTQKKVYLAKKDDYELTCTASQKKNTOFHKNKNOIK 60
Qy 61 ILGIQGSFLTKGPSKLSDRADSRKSLMDQGCFSMTIKNLKIEDSDTYICEVKNKEEVEL 120
Db 61 ILGIQGSFLTKGPSKLSDRADSRKSLMDQGCFSMTIKNLKIEDSDTYICEVKNKEEVEL 120
Qy 121 LVFGLTANSPTHLQGSGLTTLTSPSSPSVQCRSPRGKNIQCGKTLVSQLELDQSG 180
Db 121 LVFGLTANSPTHLQGSGLTTLTSPSSPSVQCRSPRGKNIQCGKTLVSQLELDQSG 180
Qy 181 TWCTCTVLONOKKVEFKIDIV 200
Db 181 TWCTCTVSODKTVKFKIDIV 200

RESULT 5
CD4_MACMU STANDARD; PRT: 458 AA.
ID CD4_MACMU
AC P16003; Q29617;
DT 01-APR-1990 (Rel. 14, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
CD4/Leu-3).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Camerini D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymocytes;
RA Hashimoto O., Tatsumi M.;
RT "Molecular cloning and expression of macaque CD4s.";
RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.E., Disotell T.R.;

```

```

RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeys (Primates: Papionini).";
RL Mol. Biol. Evol. 15:892-900(1998).
CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC - SUBUNIT: Associates with p56-lck (By similarity).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31134; AAA6838.1; -
DR EMBL; D63347; BAA09670.1; -
DR EMBL; X73326; CAAS1752.1; -
DR EMBL; AF057385; AAC25129.1; -
DR HSSP; P01730; 1MBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV_1.
DR PROSITE; PS50835; IG_LIKE_1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458 T-CELL SURFACE GLYCOPROTEIN CD4.
FT DOMAIN 26 396 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 397 418 POTENTIAL.
FT DOMAIN 419 458 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 125 IG-LIKE V-TYPE.
FT DOMAIN 126 203 IG-LIKE C2-TYPE 1.
FT DOMAIN 204 317 IG-LIKE C2-TYPE 2.
FT DOMAIN 318 374 IG-LIKE C2-TYPE 3.
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (BY SIMILARITY).
FT DISULFID 325 325 N-LINKED (GLCNAC. .) (BY SIMILARITY).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 328 370 S-palmitoyl cysteine (By similarity).
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 N -> T (IN REF. 1).
FT CONFLICT 42 42 L -> S (IN REF. 3).
FT CONFLICT 62 62 L -> S (IN REF. 2).
FT CONFLICT 67 67 I -> L (IN REF. 2).
FT CONFLICT 169 169 K -> N (IN REF. 3).
FT CONFLICT 191 191 S -> P (IN REF. 2).
FT CONFLICT 248 248 R -> Q (IN REF. 3).
FT CONFLICT 265 265 A -> T (IN REF. 2).
FT CONFLICT 349 349
SQ SEQUENCE 458 AA; 50884 MW; 8BB80339FAFC808 CRC64;

Query Match 68.6%; Score 904; DB 1; Length 458;
Best Local Similarity 87.0%; Pred. No. 1,3e-61;
Matches 174; Conservative 13; Mismatches 13; Indels 0; Gaps 0;

```

```

QY 61 IIGNOSFLTKGPKUNDRADSRSLMDGNFLLIKNLKIDSDPYICEVEDQKEVOL 120
DB 61 IIGIOFLTKGPKSKSDRSDRSKSLMDQGCFSMIITKNKIDSDPYICEVEKKEVOL 120
QY 121 LVFGLTANSDTHLQOSLTTLTLESPGSSPVOCSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLQOSLTTLTLESPGSSPVOCSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTVLTQNKYVERKIDIV 200
DB 181 TWCTVLTQNKYVERKIDIV 200

RESULT 6
CD4_MACNE STANDARD; PRT; 458 AA.
ID CD4_MACNE
AC 008340; P79196;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUN-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3).
GN CD4.
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsuami M.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RA TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Romsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
CC - FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC - SUBUNIT: Associates with p56-lck (By similarity).
CC - SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC - SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63346; BAA09670.1; -
DR EMBL; X73325; CAAS1751.1; -
DR HSSP; P01730; 1MBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.

```

DR	SMART: SM00406; IGV: 1.	
DR	PROSITE: PS00835; IG_LIKE: 1. <td></td>	
KW	Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell; <td></td>	
KW	Immune response; Repeat; Signal; Lipoprotein; Palmitate. <td></td>	
FT	SIGNAL	1 25
FT	CHAIN	26 458
FT	DOMAIN	26 396
FT	TRANSMEM	397 418
FT	DOMAIN	419 458
FT	DOMAIN	26 125
FT	DOMAIN	126 203
FT	DOMAIN	204 317
FT	DOMAIN	318 374
FT	CARBOHYD	42 42
FT	CARBOHYD	296 296
FT	CARBOHYD	325 325
FT	DISULFID	41 109
FT	DISULFID	155 184
FT	DISULFID	328 370
FT	LIPID	419 419
FT	LIPID	422 422
FT	CONFLICT	57 57
FT	CONFLICT	91 91
FT	CONFLICT	105 105
FT	CONFLICT	113 113
FT	CONFLICT	302 302
FT	CONFLICT	349 349
SO	SEQUENCE	458 AA; 50905 MW; 751A9BA2CB3BE16 CRC64;
Query Match		
Best Local Similarity		
Matches 174; Conservative 14; Mismatches 12; Indels 0; Gaps 0;		
QY	1 MNRGVPFHLHLVLTALLPATQGNKVLGKGGDVTETCTASQKSIQFHWKNSNOIK	60
Db	1 MNRGVPFHLHLVLTALLPATQGNKVLGKGGDVTETCTASQKSIQFHWKNSNOIK	60
QY	61 ILGNQSFELTKGPSKLNDRADSRSLMDQGNFLLIKLKIEDSDTYICEVEDQKEEYQL	120
Db	61 ILGNQSFELTKGPSKLNDRADSRSLMDQGNFLLIKLKIEDSDTYICEVEDQKEEYQL	120
QY	121 LVFGITANSPTHLLOGSGSLTLTLESPSSSPVQCRSPRGKTIQSGKTLISVQLELDQSG	180
Db	121 LVFGITANSPTHLLOGSGSLTLTLESPSSSPVQCRSPRGKTIQSGKTLISVQLELDQSG	180
QY	181 TWTCVTVLONOKKVEFKIDIV 200	
Db	181 TWTCVTVLONOKKVEFKIDIV 200	
RESULT 7		
CD4_CERAE . STANDARD: PRT: 458 AA.		
AC	Q08338; 002805; 077593; Q28217;	
DT	01-FEB-1995 (Rel. 31, Created)	
DT	15-MAR-2004 (Rel. 43, Last sequence update)	
DT	15-MAR-2004 (Rel. 43, Last annotation update)	
DE	T-cell surface glycoprotein CD4 precursor [T-cell surface antigen	
DE	T4/Leu-3].	
GN	CD4.	
OS	Cercopithecus aethiops (Green monkey) (Grivet).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;	
OC	Cercopitheidae; Cercopitheidae.	
OX	NCBI_TaxID=9534;	
RN	(1)	
RP	SEQUENCE FROM N.A.	
RA	Haehimoto O.; Tatsumi M.;	
RT	"Molecular cloning and expression of african green monkey CD4.";	
RL	Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.	
RP	(2)	
RP	SEQUENCE OF 28-424 FROM N.A.	
NC	TISSUE=Blood;	

```

RX MEDLINE=93049640; PubMed=1425921.
RA Pomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus." ;
RL Eur. J. Immunol. 22:2973-2981(1992).
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Periphereal blood;
RX MEDLINE=98011879; PubMed=9379478;
RA Pomsgaard A., Mellier-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corber S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes." ;
RL J. Med. Primatol. 26:120-128(1997).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.E., Disotell T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeye (Primates: Papionini)." ;
RL MoJ. Biol. Evol. 15:892-900(1998).
CC -I- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -I- SUBUNIT: Associates with p56-lck (by similarity).
CC -I- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -I- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; D86589; BAI1312.1; -.
DR DR EMBL; X73322; CA51748.1; -.
DR DR EMBL; AF001226; AAB60873.1; -.
DR DR EMBL; AF001228; AAB60875.1; -.
DR DR EMBL; AF057380; AAC5124.1; -.
DR HSSP; P01730; IWOQ.
DR DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR DR GO; GO:0015020; F:coreceptor activity; ISS.
DR DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR DR GO; GO:0006955; P:immune response; ISS.
DR DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR DR GO; GO:0045058; P:T-cell selection; ISS.
DR DR GO; GO:0007159; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPRO000973; CD4 TCAG.
DR InterPro; IPRO07110; Ig_Like.
DR InterPro; IPRO03596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PRO0692; CD4TCANTIGEN.
DR SMART; SMO0406; IGv_1.
DR PROSITE; PS50835; IG_LIKE_1.
KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipopeptide; Palmitate.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSSEM 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 482
FT CARBOHYD 281 281
FT CARBOHYD 296 296
FT CARBOHYD 325 325

```

```

FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 328 370 BY SIMILARITY.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 46 46 K -> N (IN REF. 2 AND 3; AAB60875).
FT CONFLICT 59 59 I -> T (IN REF. 3; AAB60873).
FT CONFLICT 115 115 K -> E (IN REF. 1).
FT CONFLICT 155 155 G -> V (IN REF. 3; AAB60873 AND 4).
FT CONFLICT 165 165 M -> L (IN REF. 3; AAB60873).
FT CONFLICT 200 200 F -> L (IN REF. 3; AAB60873).
FT CONFLICT 227 227 K -> E (IN REF. 3; AAB60873).
FT CONFLICT 271 271 N -> H (IN REF. 3; AAB60873).
FT CONFLICT 281 281 N -> H (IN REF. 3; AAB60873).
SQ SEQUENCE 458 AA; 51158 MW; FC523D2EDD1F72E7 CRC64;

Query Match 67.2%; Score 885; DB 1; Length 458;
Best Local Similarity 85.5%; Pred. No. 3,7e-60;
Matches 171; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

QY 1 MNRGVPRLHLLVQLALPAATQGNKVVLGKGGDTVELCTASQKKSIOFHKNNOIK 60
DB 1 MNRGVPRLHLLVQLALPAATQGNKVVLGKGGDTVELCTASQKKSIOFHKNNOIK 60
QY 61 ILGQGSFLTKGSKLNDRAISRSLMDQGNPLIINKLIEDSDTYICEVEDQKEVOL 120
DB 61 ILGQGSFLTKGSKLNDRAISRSLMDQGNPLIINKLIEDSDTYICEVEDQKEVOL 120
QY 121 LVFGLTNSDTHLLQGGSLTLTLSPGSSPVQCRSPRGKNTQGGKTLVSQLELDSG 180
DB 121 LVFGLTNSDTHLLQGGSLTLTLSPGSSPVQCRSPRGKNTQGGKTLVSQLELDSG 180
QY 181 TWICTVONOKKVEFKIDIV 200
DB 181 TWICTVONOKKVEFKIDIV 200
QY 181 TWICTVONOKKVEFKIDIV 200
DB 181 TWICTVONOKKVEFKIDIV 200

RESULT 8
CD4_CERTO STANDARD; PRT; 397 AA.
ID CD4_CERTO
AC Q08336;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)
DE (Fragment).
GN CD4.
OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OC NCBI_TaxID=9531;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood.
RX MEDLINE=39049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus."
RT J. Immunol. 22:2973-2981 (1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announc/

```

```

CC or send an email to license@sb-sib.ch).
CC EMBL: X73328; CAA51754.1; -.
CC EMBL: X73327; CAA51753.1; -.
CC HSSP; P01730; 1MIQ.
DR GO; GO:0042101; C: T-cell receptor complex; ISS.
DR GO; GO:0045026; F: coreceptor activity; ISS.
DR GO; GO:0042289; F: MHC class II protein binding; ISS.
DR GO; GO:0005955; P: immune response; ISS.
DR GO; GO:0045086; P: positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P: T-cell differentiation; ISS.
DR GO; GO:0045058; P: T-cell selection; ISS.
DR GO; GO:0007169; P: transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR009973; CD4_TcRg.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Lipoprotein; Palmitate.
FT NON_TER 1
FT DOMAIN <1 369 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 370 391 POTENTIAL.
FT DOMAIN 392 >397 CYTOPLASMIC (POTENTIAL).
FT DOMAIN <1 98 IG-LIKE V-TYPE.
FT DOMAIN 99 176 IG-LIKE C2-TYPE 1.
FT DOMAIN 177 290 IG-LIKE C2-TYPE 2.
FT DOMAIN 291 347 IG-LIKE C2-TYPE 3.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 14 82 BY SIMILARITY.
FT DISULFID 128 157 BY SIMILARITY.
FT LIPID 301 343 S-palmitoyl cysteine (By similarity).
FT DISULFID 392 392 S-palmitoyl cysteine (By similarity).
FT LIPID 395 395 S-palmitoyl cysteine (By similarity).
FT VARIANT 20 20 MISSING.
FT VARIANT 43 43 T -> I.
FT VARIANT 86 86 N -> D.
FT VARIANT 96 96 F -> L.
FT VARIANT 173 173 V -> M.
FT VARIANT 316 316 R -> K.
FT NON_TER 397
SQ SEQUENCE 397 AA; 43926 MW; 8660B636D2DB38A7 CRC64;

Query Match 59.5%; Score 784; DB 1; Length 397;
Best Local Similarity 87.3%; Pred. No. 1.5e-52;
Matches 151; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 28 VVLGKGGDTVELCTASQKKSIOFHKNNOIKILQGGSLTKGSKLNDRAISRSLW 87
DB 1 VVLGKGGDTVELCTASQKKSIOFHKNNOIKILQGGSLTKGSKLNDRAISRSLW 87
QY 88 DQGNFPLIINKLIEDSDTYICEVEDQKEVOLVGLTNSDTHLLQGGSLTLTLSP 147
DB 61 DQGNFPLIINKLIEDSDTYICEVEDQKEVOLVGLTNSDTHLLQGGSLTLTLSP 120
QY 148 GSSPSVQCRSPRGKNTQGGKTLVSQLELDSGTTWCTVLOKVEFKIDIV 200
DB 121 GSSPSVQCRSPRGKNTQGGKTLVSQLELDSGTTWCTVLOKVEFKIDIV 173

RESULT 9
CD4_ERYPA STANDARD; PRT; 397 AA.
ID CD4_ERYPA
AC Q08339;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)
DE (Fragment).

```

```

SN      CD4.
NS      Erythrocebus patas (Red guenon) (Husgar).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC      Cercopithecinae; Erythrocebus.
OX      NCBI_TaxId=9538;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Blood;
RX      MEDLINE=93049640; PubMed=14259211;
RA      Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT      Cloning and sequences of primate CD4 molecules: diversity of the
RT      cellular receptor for simian immunodeficiency virus/human
RT      immunodeficiency virus.";
RL      Eur. J. Immunol. 22:2973-2981(1992).
CC      -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC      receptor interaction. May regulate T-cell activation.
CC      -1- SUBUNIT: Associates with p56-lck (By similarity).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X73324; CA51750.1; -.
DR      HSSP; P01730; 1W1Q.
DR      GO; GO:0042101; C:T-cell receptor complex; ISS.
DR      GO; GO:0015026; F:coreceptor activity; ISS.
DR      GO; GO:0042289; F:MHC class II protein binding; ISS.
DR      GO; GO:0006955; P:immune response; ISS.
DR      GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR      GO; GO:0030217; P:T-cell differentiation; ISS.
DR      GO; GO:0045058; P:T-cell selection; ISS.
DR      GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR      InterPro; IPR000973; Ctd_TcOAg.
DR      InterPro; IPR002110; Ig-Like.
DR      InterPro; IPR003596; Ig_V.
DR      Pfam; PF00047; Ig; 2.
DR      PRINTS; PR00692; CD4TCANTIGEN.
DR      SMART; SM00406; IGV; 1.
DR      PROSITE; PS50835; IG_LIKE; 1.
KM      Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KM      Immune response; Repeat; Lipoprotein; Palmitate.
FT      NON_TER 1 1
FT      DOMAIN <1 369 EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM 370 391 POTENTIAL.
FT      DOMAIN 392 >397 CYTOPLASMIC (POTENTIAL).
FT      DOMAIN <1 98 IG-LIKE V-TYPE.
FT      DOMAIN 99 176 IG-LIKE C2-TYPE 1.
FT      DOMAIN 177 290 IG-LIKE C2-TYPE 2.
FT      DOMAIN 291 347 IG-LIKE C2-TYPE 3.
FT      CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT      CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT      DISULFID 14 82 BY SIMILARITY.
FT      DISULFID 128 157 BY SIMILARITY.
FT      DISULFID 301 343 BY SIMILARITY.
FT      LIPID 392 392 S-palmitoyl cysteine (By similarity).
FT      LIPID 395 395 S-palmitoyl cysteine (By similarity).
FT      NON_TER 397 397
SQ      SEQUENCE 397 AA; 44081 MW; 67887397A6B7EA4F CRC64;
Query Match 59.5%; Score 783; DB 1; Length 397;
Beat Local Similarity 86.7%; Pred. NO. 1.8e-52;
Matches 150; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

```

```

Db      1 VVLGKGQVTELTNCASQKTTTGTHMKNSNMOKMLKLGQSLTGTGSPKLAIDRDSRSLSW 60
Qy      88 DQGNFLLIKNLKIEDSDTYICEVEDOKEEVOLLVFGLTANSDFHLLQGSLTTLTLESP 147
Db      61 DQGFPMIIRKLIKIEDSETYICEVEDKEEVELLVFGLTANSDFHLLQGSLTTLTLESP 120
Qy      148 GSSPVCORSRGKNIQGKTLTSLQFLDQSGTWCTCTVQLNQCKVEFKIDIV 200
Db      121 GSSPVCORSRGKNIQGKTLTSLQFLDQSGTWCTCTVSDQNTVEFKIDIV 173

RESULT 10
CD4_SAISC
ID      CD4_SAISC      STANDARD;      PRT;      457 AA.
AC      Q29037;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
       T4/Lew-3).
GN      CD4.
OS      Saimiri leureus (Common squirrel monkey).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX      NCBI_TaxID=95211;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Tatemura M., Hashimoto O.;
RL      Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC      receptor interaction. May regulate T-cell activation.
CC      -1- SUBUNIT: Associates with p56-lck (by similarity).
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
CC      EMBL; D86588; BAAL3131.1; -.
DR      HSSP; P01730; 1WBR.
DR      GO; GO:0042101; C:T-cell receptor complex; ISS.
DR      GO; GO:0015026; F:coreceptor activity; ISS.
DR      GO; GO:0042289; F:MHC class II protein binding; ISS.
DR      GO; GO:0006955; P:immune response; ISS.
DR      GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. ; ISS.
DR      GO; GO:0030217; P:T-cell differentiation; ISS.
DR      GO; GO:0045058; P:T-cell selection; ISS.
DR      GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . ; ISS.
DR      InterPro; IPR000973; CD4 TcAg.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003596; Ig_v.
DR      Pfam; PF00047; Ig_2.
DR      PRINTS; PR00692; CD4TCANTIGEN.
DR      SMART; SM00406; IGV_1.
DR      PROSITE; PS50835; IG_LIKE; 1.
KW      Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
       Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT      SIGNAL 1 25
FT      CHAIN 26 457
FT      DOMAIN 26 395
FT      DOMAIN 396 417
FT      TRANSMEM 418 457
FT      DOMAIN 126 125
FT      DOMAIN 126 102
FT      DOMAIN 203 316
FT      DOMAIN 317 373
FT      CARBOHYD 254 254
       N-LINKED (GLCNAC. .) (POTENTIAL).

```



```

FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 41 109 BY SIMILARITY.
FT DISULFID 155 184 BY SIMILARITY.
FT DISULFID 327 369 BY SIMILARITY.
FT LIPID 418 418 S-palmitoyl cysteine (By similarity).
FT LIPID 421 421 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 457 AA; 50871 MW; 57EB6344005A015 CRC64;

Query Match 56.0%; Score 737.5; DB 1; Length 457;
Best Local Similarity 71.5%; Pred. No. 6,2e-49;
Matches 143; Conservative 24; Mismatches 32; Indels 1; Gaps 1;

QY 1 MNRGVPFRHLLVQLALPAATQGNKVLGKGGDYELTCTASQKSIQPHKNSNQIK 60
DB 1 MNGGIPFRHLLVQLALPAATQGNKVLGKGGDYELTCTASQKSIQPHKNSNQIK 60
QY 61 ILGNQGSFLLTKGPKSKINDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLLTKGPKSKINDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVEDQKEEVOL 120
QY 61 ILGVQNFVTRGSGSKLTRIDSKSSMDRSGFPLIDARIEDSETTYICEVESKEEVEL 120
DB 121 LVEGLTANSDTHLLQGSLLTLLSPGSSPSVQCSPRGKNIQGGKTLSSVQLQLDQSG 180
QY 121 QVGLTANPDTHLLQGSLLTLLSPGSSPSVQCTSPRGKIRGRKTLSSVQLGIPDSG 180
DB 121 QVGLTANPDTHLLQGSLLTLLSPGSSPSVQCTSPRGKIRGRKTLSSVQLGIPDSG 180
QY 181 TWTCVQLQNKVQKFEKIDIV 200
DB 181 TWTCVQLQNKVQKFEKIDIV 200
QY 181 TWTCVQLQNKVQKFEKIDIV 199
DB 181 TWTCVQLQNKVQKFEKIDIV 199

RESULT 11
ID CD4_CANFA STANDARD; PRT; 463 AA.
AC P33705;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Lew-3).
GN CD4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE OF 13-463 FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
RX MEDLINE=93192324; PubMed=7916632;
RA Milde K.F., Conner G.E., Minz D.H., Alejandro R.;
RT "Primary structure of the canine CD4 antigen.";
RL Biochim. Biophys. Acta 1172:315-318(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Beagle; TISSUE=Thymus;
RX MEDLINE=94378217; PubMed=8091416;
RA Gorman S.D., Frewin M.R., Cobbold S.P., Waldmann H.;
RT "Isolation and expression of cDNA encoding the canine CD4 and CD8
alpha antigens.";
RL Tissue Antigens 43:184-188(1994).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in macrophages and a subset of
T lymphocytes.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entries requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L06130; AAB02295.1; -.
CC EMBL; X68565; -, NOT_ANNOTATED_CDS.
DR HSSP; P01730; 1MBR.
DR GO; GO:0042101; C-T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR000973; CD4_TCRG.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_3.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KW Immune response; Repeat; Signal; Lipoprotein; Palmitate.
FT SIGNAL 1 24
FT CHAIN 25 463
FT DOMAIN 25 401
FT TRANSMEM 402 423
FT DOMAIN 424 463
FT DOMAIN 26 124
FT DOMAIN 125 211
FT DOMAIN 212 321
FT DOMAIN 322 378
FT DISULFID 41 109
FT DISULFID 332 374
FT LIPID 424 424
FT LIPID 427 427
FT CARBOHYD 123 123
FT CARBOHYD 168 168
FT CARBOHYD 176 176
FT CARBOHYD 324 324
FT CARBOHYD 329 329
FT CARBOHYD 389 389
SQ SEQUENCE 463 AA; 51639 MW; 95805170CB44A833 CRC64;

Query Match 45.4%; Score 598; DB 1; Length 463;
Best Local Similarity 55.5%; Pred. No. 2.7e-38;
Matches 116; Conservative 38; Mismatches 45; Indels 10; Gaps 2;

QY 1 MNRGVPFRHLLVQLALPAATQGNKVLGKGGDYELTCTASQKSIQPHKNSNQIK 60
DB 1 MNGGIPFRHLLVQLALPAATQGNKVLGKGGDYELTCTASQKSIQPHKNSNQIK 60
QY 61 ILGNQGSFLLTKGPKSKINDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGNQGSFLLTKGPKSKINDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVEGLTANSDTHLLQGSLLTLLSPGSSPSVQCSPRGKNIQGGKTLSSVQLQLDQSG 171
DB 121 LVEGLTANSDTHLLQGSLLTLLSPGSSPSVQCSPRGKNIQGGKTLSSVQLQLDQSG 171
QY 121 LVEGLTANSDTHLLQGSLLTLLSPGSSPSVQCSPRGKNIQGGKTLSSVQLQLDQSG 179
DB 121 LVEGLTANSDTHLLQGSLLTLLSPGSSPSVQCSPRGKNIQGGKTLSSVQLQLDQSG 179
QY 172 SOLELDGSGTWCTVQLQNKVQKFEKIDIV 200
DB 172 SOLELDGSGTWCTVQLQNKVQKFEKIDIV 200
QY 180 SWPELDGSGTWCTVQLQNKVQKFEKIDIV 208
DB 180 SWPELDGSGTWCTVQLQNKVQKFEKIDIV 208

RESULT 12
ID CD4_RABIT STANDARD; PRT; 459 AA.
AC P46630;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen

```

DE T4/Leu-3).

GN CD4.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92390370; PubMed=1518821;

RA Hague B.F., Sawadkoseol S., Brown T.J., Lee K., Recker D.P.,

RA Kindt T.J.;

RT "CD4 and its role in infection of rabbit cell lines by human

RT immunodeficiency virus type 1."

RL Proc. Natl. Acad. Sci. U.S.A. 89:7963-7967(1992).

CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell

CC receptor interaction. May regulate T-cell activation.

CC -1- SUBUNIT: Associates with p56-lck (By similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; M92840; AAA31198.1; -

DR PIR; A46254; A46254.

DR HSP; P01730; 1MBR.

DR GO; GO:0042101; C:T-cell receptor complex; ISS.

DR GO; GO:0015026; F:coreceptor activity; ISS.

DR GO; GO:0042289; F:MHC class II protein binding; ISS.

DR GO; GO:0006955; P:immune response; ISS.

DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.

DR GO; GO:0030217; P:T-cell differentiation; ISS.

DR GO; GO:0045058; P:T-cell selection; ISS.

DR GO; GO:0007159; P:transmembrane receptor protein tyrosine kin. . .; ISS.

DR InterPro; IPR000973; CD4_TcAg.

DR InterPro; IPR007110; Ig_Like.

DR Pfam; PF00047; Ig_2.

DR PRINTS; PR00692; CD4TCANTIGEN.

DR SMART; SM00406; IG_1.

DR PROSITE; PS00835; IG_LIKE; 1.

KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;

CC Immune response; Repeat; Signal; Lipoprotein; Palmitate.

FT SIGNAL 1 25

FT CHAIN 26 459

FT DOMAIN 26 396

FT TRANSSEM 397 419

FT DOMAIN 420 459

FT DOMAIN 26 129

FT DOMAIN 130 208

FT DOMAIN 209 318

FT DOMAIN 319 374

FT CARBOHYD 299 299

FT DISULFID 41 113

FT DISULFID 329 370

FT LIPID 420 423

FT LIPID 423 423

SEQ SEQUENCE 459 AA; 50886 MW; B32311CBDA0013D CRC64;

Query Match 45.3%; Score 596.5; DB 1; Length 459;

Best Local Similarity 57.9%; Pred. No. 3.4e-38;

Matches 124; Conservative 35; Mismatches 46; Indels 9; Gaps 3;

OY 1 MNRGVPRRLHLVQLALPAATGKRVKLGKGDVLELTCTASQKKSIOFHWNKSNQIK 60

DB 1 MNRRIYFOCLLVPLALPPAATGKRVKRGKAGAIIVELPSSQSKNSVENVWGHANQVK 60

OY 61 ILNGQ-----SFLTKGPSKINDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDOKE 116

DB 61 ILNGQSSSSSFWLKGKSPPLSNRVESKKNMMQDSFFLVIKDLRMDSGTYICEVGRK 120

OY 117 EVOLVGLTANSDTHLLQGSLLTLTSPGSSPSVQCSPRKNIQGGKTLSSVQLEL 176

DB 121 EVELLVRLTANPTRLHQQSLTLTLEGSPSVQVMSPEKKIETGPTCSMPKRL 180

OY 177 ODGCTWCTV-LDNQKVEFEKIDV----PRAS 205

DB 181 QDSGTWSCHLSFDQNTLELDIKITLVGFPRASA 214

RESULT 13

ID CD4_RAT STANDARD; PRT; 457 AA.

AC P05540;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen

DE T4/Leu-3) (M3/25 antigen).

GN CD4.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RX MEDLINE=8717535; PubMed=1104900;

RA Clark S.J., Jefferies W.A., Barclay A.N., Gagnon J., Williams A.F.;

RT "Peptide and nucleotide sequences of rat CD4 (M3/25) antigen:

RT evidence for derivation from a structure with four

RT immunoglobulin-related domains."

RL Proc. Natl. Acad. Sci. U.S.A. 84:1649-1653(1987).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 210-393.

RX MEDLINE=93262437; PubMed=8493535.

RA Brady R.L., Dodson E.T., Dodson G.G., Lange G., Davis S.J.,

RA Williams A.F., Barclay A.N.;

RT "Crystal structure of domains 3 and 4 of rat CD4: relation to the

RT NH2-terminal domains."

RL Science 260:979-983(1993).

CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell

CC receptor interaction. May regulate T-cell activation.

CC -1- SUBUNIT: Associates with p56-lck (By similarity).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; M15768; AAA40901.1; -

DR PIR; A27449; A27449.

DR PDB; 1CID; 15-JUL-93.

DR GlycoStated; P05540; -

DR GO; GO:0042101; C:T-cell receptor complex; ISS.

DR GO; GO:0015026; F:coreceptor activity; ISS.

DR GO; GO:0042289; F:MHC class II protein binding; ISS.

DR GO; GO:0006955; P:immune response; ISS.

DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.

DR GO; GO:0030217; P:T-cell differentiation; ISS.

DR GO; GO:0045058; P:T-cell selection; ISS.

DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.

DR InterPro; IPR000973; CD4_TcAg.

DR InterPro; IPR007110; Ig_Like.

DR InterPro; IPR003599; Ig.

```

DR pfam: PF00047; ig; 2.
DR PRINTS; PRO0692; CD4TANTIGEN.
DR SMART; SM00409; ig; 2.
DR PROSITE; PSS00835; IG_LIKE; 1.
KM Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
KM Immune response; Repeat; Signal; Lipoprotein; Palmitate; 3D-structure.
FT SIGNAL 1 27
FT CHAIN 28 457
FT DOMAIN 28 394
FT TRANSLEM 395 417
FT DOMAIN 418 457
FT DOMAIN 28 127
FT DOMAIN 128 206
FT DOMAIN 207 316
FT DOMAIN 317 374
FT CARBOHYD 186 186
FT CARBOHYD 297 297
FT CARBOHYD 392 392
FT DISULFID 43 111
FT DISULFID 158 187
FT LIPID 328 370
FT LIPID 418 418
FT LIPID 421 421
FT STRAND 213 217
FT TURN 218 219
FT STRAND 222 225
FT STRAND 235 243
FT STRAND 252 258
FT TURN 259 260
FT STRAND 261 265
FT STRAND 274 275
FT STRAND 278 278
FT TURN 279 280
FT STRAND 282 285
FT HELIX 290 292
FT STRAND 294 301
FT STRAND 306 319
FT STRAND 325 331
FT STRAND 338 344
FT TURN 345 346
FT STRAND 349 353
FT STRAND 357 361
FT STRAND 367 374
FT TURN 375 376
FT STRAND 377 385
SQ SEQUENCE 457 AA; 51437 MW; 477B157D30954C1 CRC64;

Query Match 37.2%; Score 490.5; DB 1; Length 457;
Best Local Similarity 48.8%; Pred. No. 4e-30;
Matches 102; Conservative 35; Mismatches 67; Indels 5; Gaps 3;

QY 1 MNRGVPRRHL-LIVLQALIPATGNGKVVGLGKGDVTBELTCTASQKSIQPHWKNQ 58
DB 1 MCGFSPRRHLPLLLLOLSKLLVVTQGTQVTLGREGSAELPCESTSRBSAFMKSSDQ 60
QY 59 IKILNGSGSLTQSPKSLNDRADSRSLMOGNEPLLTKLKIEDSPTYICEVDQEEV 118
DB 61 KTIIGVKNKLKIKSLSLYSFDSRKNAMKSGPPLINKLRMEDSOTVYCELENKKEEV 120
QY 119 QLVFGLTANSDELTLQGSITLTLES-PPGSSPSVOCRRPRGNIGGKTLVSQLELQ 177
DB 121 ELWFRRTFNPGRITLLQGSITLTLIDNPKVSDDPICRHKSSIVVDSKAFTHSLRIQ 180
QY 178 DSGTWTCTVLONQKKEVF--KIDIVPRAS 204
DB 181 DSGIWNCTVTLNQRKHSFDMKLSVLCFAS 209

RESULT 14
ID CD4 MOUSE STANDARD; PRT; 457 AA.
AC P06332;
DT 01-JAN-1988 (Rel. 06, Created)

```

```

DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
DE T4/Leu-3) (T-cell differentiation antigen L3T4).
GN CD4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=87018845; PubMed=3094146;
RT Tourville B., Gorman S.D., Field E.H., Hunkapiller T., Parnes J.R.;
RT "Isolation and sequence of L3T4 complementary DNA clones: expression
RT in T cells and brain.";
RL Science 234:610-614(1986).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=87115821; PubMed=3027575;
RT Littman D.R., Gettner S.N.;
RT "Unusual intron in the immunoglobulin domain of the newly isolated
RT murine CD4 (L3T4) gene.";
RL Nature 325:453-455(1987).
RN (3)
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=88152875; PubMed=3326818;
RA Parnes J.R., Hunkapiller T.;
RT "L3T4 and the immunoglobulin gene superfamily: new relationships
RT between the immune system and the nervous system.";
RL Immunol. Rev. 100:109-127(1987).
RN (4)
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Brain;
RX MEDLINE=88041159; PubMed=2823269;
RA Gorman S.D., Tourville B., Parnes J.R.;
RT "Structure of the mouse gene encoding CD4 and an unusual transcript
RT in brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7644-7648(1987).
RN (5)
RP SEQUENCE FROM N.A.
RX MEDLINE=98112780; PubMed=9445485;
RA Ansari-Lari M.A., Oeltjen J.C., Schwartz S., Zhang Z., Muzny D.M.,
RA Lu J., Gorrell J.H., Chinnault A.C., Belmont J.W., Miller W.,
RA Gibbs R.A.;
RT "Comparative sequence analysis of a gene-rich cluster at human
RT chromosome 12p13 and its syntenic region in mouse chromosome 6.";
RL Genome Res. 8:29-40(1998).
RN (6)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX STRAIN=C57BL/6J; TISSUE=Mammary gland;
RX MEDLINE=22368257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares W.B., Bonaldo W.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Locantello N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]

```

RP SEQUENCE OF 27-43.
 RX MEDLINE=8616694; PubMed=3082751;
 RA Claesson B.J., Teagarator J., Kiersbaum L., Maddox J., McKay C.R.,
 RA Brandon M., McKenzie I.F.C., Walker I.D.;
 RT "The LIT4 antigen in mouse and the sheep equivalent are
 RT immunoglobulin-like.";
 RL Immunogenetics 23:129-132(1986).
 RN [8]
 RP DISULFIDE BONDS.
 RA MEDLINE=86233454; PubMed=3086886;
 RA Claesson B.J., Teagarator J., McKenzie I.F.C., Walker I.D.;
 RT "Partial primary structure of the T4 antigens of mouse and sheep:
 RT assignment of intrachain disulfide bonds.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:4499-4503(1986).
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
 CC receptor interaction. May regulate T-cell activation.
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P06332-1; Sequence=Displayed;
 CC Name=2; Synonyms=Brain-specific;
 CC IsoId=P06332-2; Sequence=VSP_002489;
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M36850; AAA39401.1; -;
 DR EMBL; M13816; AAA37267.1; -;
 DR EMBL; X04836; CA28539.1; -;
 DR EMBL; M36851; AAA39402.1; -;
 DR EMBL; M17080; AAA37403.1; -;
 DR EMBL; M17078; AAA37403.1; JOINED.
 DR EMBL; M17079; AAA37403.1; JOINED.
 DR EMBL; AC002397; AAC36010.1; -;
 DR EMBL; BC039137; AAC39137.1; -;
 DR PIR; A02110; RWMST4.
 DR HSSP; P01730; IWR.
 DR MGD; MGI:88335; Cd4.
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.
 DR GO; GO:0015026; F:coreceptor activity; ISS.
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.
 DR GO; GO:0006955; P:immune response; ISS.
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
 DR GO; GO:0030217; P:T-cell differentiation; ISS.
 DR GO; GO:0045058; P:T-cell selection; ISS.
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
 DR InterPro; IPR000973; CD4 TCRg.
 DR InterPro; IPR007110; Ig-Tlike.
 DR InterPro; IPR003596; Ig_V.
 DR Pfam; PF00047; Ig_2.
 DR PRINTS; PRO0692; CD4TCANTIGEN.
 DR SMART; SMO0406; IGV_1.
 DR PROSITE; PS50835; IG-LIKE_1.
 KW Immunoglobulin domain; Transmembrane; Glycoprotein; T-cell;
 KW Immune response; Repeat; Signal; Lipoprotein; Palmitate;
 KW Alternative splicing.
 FT SIGNAL 1 26
 FT CHAIN 27 457 T-CELL SURFACE GLYCOPROTEIN CD4.
 FT DOMAIN 27 394 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 395 417 POTENTIAL.
 FT DOMAIN 418 457 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 27 128 IG-LIKE V-TYPE.
 FT DOMAIN 129 207 IG-LIKE C2-TYPE 1.
 FT DOMAIN 208 317 IG-LIKE C2-TYPE 2.

FT DOMAIN 318 374 IG-LIKE C2-TYPE 3.
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 42 112
 FT DISULFID 159 188
 FT DISULFID 328 370
 FT LIPID 418 418 S-palmitoyl cysteine (By similarity).
 FT LIPID 421 421 S-palmitoyl cysteine (By similarity).
 FT VAAPPLIC 1 240 Missing (in isoform 2).
 SQ SEQUENCE 457 AA; 51296 MW; 1B1DA7527CB00F33 CRC64;
 FT FT
 Query Match 36.1%; Score 475; DB 1; Length 457;
 Best Local Similarity 52.8%; Pred. No. 6,1e-25;
 Matches 104; Conservative 32; Mismatches 55; Indels 6; Gaps 5;
 Oy 1 MNRGVPRH-LLVLTAL-PAATQGNKVVYGGKGDVETLTCTASOKKSIQFHKNSNQI 59
 Db 1 MCRALSRLLILLQLSLAVTQKTVLKGESAEFLPSSQKKITVFTYKPSDOR 60
 Oy 60 KIILNQG-SPLYTG--PSKINDRADSRSLMDQNPPLIKNLKIEDSDTYICEVEDQKE 116
 Db 61 KIIGQHGKGVLRGSGSPSQF-DREDSKKGAWKESFPLINKLWEDSQTYICELENRKE 119
 Oy 117 EVQLVIGLTRANSTHLLQGSLLTLES-PGSSPSVQCRSPGKXIQGKITSVQLE 175
 Db 120 EVELMVKVFVSPSPSTLSLQGSLLTLTLDNSKSNPLTECKGKXKGVSGKVSMSLR 179
 Oy 176 LDGSGTWCCTVYLOQKK 192
 Db 180 VQSDPFNCTVTLDPKK 196
 RESULT 15
 CD7 HUMAN STANDARD; PRT; 240 AA.
 AC P09564;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE T-cell antigen CD7 precursor (GP40) (T-cell leukemia antigen) (TP41)
 DE (Leu-9).
 GN CD7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8811517; PubMed=3501369;
 RA Aruffo A., Seed B.;
 RT "Molecular cloning of two CD7 (T-cell leukemia antigen) cDNAs by a
 RT COS cell expression system.";
 RL EMBO J. 6:3313-3316(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9110576; PubMed=1703303;
 RA Schanberg L.B., Flenner D.E., Kurtzberg J., Haynes B.F., Kaufman R.E.;
 RT "Isolation and characterization of the genomic human CD7 gene:
 RT structural similarity with the murine Thy-1 gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:603-607(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toehiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huilyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch J., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RA [4]
 RN SEQUENCE OF 205-240 FROM N.A.
 RX MEDLINE=91267564; PubMed=1711009;
 RA Yoshikawa K., Seto M., Ueda R., Obata Y., Notoke K., Yokochi T.,
 RA Takahashi T.,
 RT "Molecular cloning of the gene coding for the human T cell
 RT differentiation antigen CD7".
 RL Immunogenetics 33:352-360(1991).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE=90063052; PubMed=2479685;
 RA Ware R.E., Scaer R.M., Dietz M.A., Starmer C.F., Palker T.J.,
 RA Haynes B.F.,
 RT "Characterization of the surface topography and putative tertiary
 RT structure of the human CD7 molecule".
 RL J. Immunol. 143:3632-3640(1989).
 RN [6]
 RN INTERACTION WITH SECTM1.
 RX MEDLINE=20119303; PubMed=10652336;
 RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.,
 RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
 RT protein".
 RL J. Biol. Chem. 275:3431-3437(2000).
 CC -|- FUNCTION: Not yet known.
 CC -|- SUBUNIT: Interacts with SECTM1.
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -|- SIMILARITY: Contains 1 immunoglobulin-like domain.
 CC -|- DATABASE: NAME=PROV; NOTE=CD guide CD7 entry;
 CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd7.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: X06180; CAA29546.1; -;
 DR EMBL: M37271; AAA51953.1; -;
 DR EMBL: BC009293; AAH09293.1; -;
 DR EMBL: BC013287; AAH13287.1; -;
 DR EMBL: D00749; BAA00646.1; -;
 DR EMBL: D00747; BAA00646.1; JOINED.
 DR EMBL: D00748; BAA00646.1; JOINED.
 DR PIR: A39016; A39016.
 DR HSSP: P01607; IREI.
 DR Genew; HGNC:1695; CD7.
 DR MIM: 186820; -;
 DR GO: GO:0016021; C:integral to membrane; TAS.
 DR GO: GO:0005624; C:membrane fraction; TAS.
 DR GO: GO:0005886; C:plasma membrane; TAS.
 DR GO: GO:0004872; F:receptor activity; TAS.
 DR GO: GO:0006955; P:immune response; TAS.
 DR GO: GO:0042110; P:T-cell activation; TAS.
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. .; NAS.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003599; IG.
 DR Pfam; PF00047; Ig; 1.

DR SMART; SM00409; IG; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW T-cell; Signal; Immune response; Antigen; Transmembrane; Glycoprotein;
 KW Immunoglobulin domain; Receptor; Lipoprotein; Palmitate; Repeat.
 FT SIGNAL 1 25
 FT CHAIN 26 240
 FT DOMAIN 26 180
 FT TRANSMEM 181 201
 FT DOMAIN 202 240
 FT DOMAIN 26 130
 FT DOMAIN 145 180
 FT DISULFID 35 142
 FT DISULFID 48 114
 FT LIPID 198 198
 FT CARBOHYD 45 45
 FT CARBOHYD 96 96
 FT REPEAT 145 153
 FT REPEAT 154 162
 FT REPEAT 163 171
 FT REPEAT 172 180
 SQ SEQUENCE 240 AA; 25409 MW; EBBCE08279552108 CRC64;
 Query Match 25.2%; Score 332; DB 1; Length 240;
 Best Local Similarity 36.4%; Pred. No. 2,2e-18;
 Matches 102; Conservative 18; Mismatches 54; Indels 106; Gaps 10;
 QY 6 PERHLVLTQLAL--LPATQGNKVLGK-----GDTVELCTASQKSIQFHWKNS 56
 DB 4 PPRLLPLLLALARGLPALAAQEVQSPHCTVVGASVITCSTG----- 52
 QY 57 NQKILGNGSFLTK-GPS-----KANDRADSRSLNDQGNFLIKN 98
 DB 53 -----GLKGIYLRQGPQPODIYYEDGVPTDRFRGRIDFGS---QDLTTTHNR 103
 QY 99 LKIEDSDTYICEVEDQKEEVQLLVFGLTANSDTHLLQGSLLTLESPPGSSPSVQCRSP 158
 DB 104 LQSLDGTGYTC-----QAIIV----- 120
 QY 159 RGNKIQGKTLVSQLELDQSGTWCTVLQONQKVEFKIDIVPRASALPAPPGSALPDP 218
 DB 121 ---NVGSGTL-VLVTEEGSQGWNRCS-----DAPPRASALPAPPGSALPDP 164
 QY 219 QTPASALPPRPSALPALAVISFLIGLVACVLAARR 258
 DB 165 QTPASALPPRPSALPALAVISFLIGLVACVLAARR 204
 RESULT 16
 CAML_RAT STANDARD; PRT; 1259 AA.
 AC Q05695;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neural cell adhesion molecule L1 precursor (N-CAM L1).
 GN L1CAM OR CAML1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=91372414; PubMed=1894011;
 RA Mura M., Kobayashi M., Abou H., Dyemura K.,
 RT "Molecular cloning of cDNA encoding the rat neural cell adhesion
 RT molecule L1. Two L1 isoforms in the cytoplasmic region are produced
 RT by differential splicing".
 RL FEBS Lett. 289:91-95(1991).
 CC -|- FUNCTION: Cell adhesion molecule with an important role in the
 CC development of the nervous system. Involved in neuron-neuron
 CC adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
 CC to axonin on neurons.

```

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q05695-1; Sequence=Displayed;
CC Name=2; Synonym=ALics;
CC IsoId=Q05695-2; Sequence=VSP_002592;
CC TISSUE SPECIFICITY: Isoform 2 is predominantly found in the brain,
CC while isoform 1 is found in the peripheral nervous system.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC L1/neurofascin/NG2CAM family.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59149; CA41860.1; -.
CC PIR; S36126; S36126.
CC HSSP; P20241; ICFB.
CC InterPro: IPR008957; FN_III-like.
CC InterPro: IPR003961; FN_III.
CC InterPro: IPR007110; IG-like.
CC InterPro: IPR003598; IG_c2.
CC Pfam; PF00041; fn3; 4.
CC Pfam; PF00047; Ig; 6.
CC SMART; SM00060; FN3; 4.
CC SMART; SM00408; IGC2; 5.
CC PROSITE; PS00835; IG_LIKE; 6.
CC Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;
CC Transmembrane; Repeat; Immunoglobulin domain; Signal;
CC Alternative splicing.
CC FT SIGNAL 1 19 BY SIMILARITY.
CC FT CHAIN 20 1259 NEURAL CELL ADHESION MOLECULE L1.
CC FT DOMAIN 20 1122 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 1123 1145 POTENTIAL.
CC FT DOMAIN 1146 1259 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 138 128 IG-LIKE C2-TYPE 1.
CC FT DOMAIN 239 125 IG-LIKE C2-TYPE 2.
CC FT DOMAIN 332 327 IG-LIKE C2-TYPE 3.
CC FT DOMAIN 424 419 IG-LIKE C2-TYPE 4.
CC FT DOMAIN 517 600 IG-LIKE C2-TYPE 5.
CC FT DOMAIN 827 896 IG-LIKE C2-TYPE 6.
CC FT DOMAIN 932 994 FIBRONECTIN TYPE-III 1.
CC FT DOMAIN 1032 1093 FIBRONECTIN TYPE-III 2.
CC FT SITE 553 555 CELL ATTACHMENT SITE (POTENTIAL).
CC FT SITE 562 564 CELL ATTACHMENT SITE (POTENTIAL).
CC FT DISULFID 57 113 BY SIMILARITY.
CC FT DISULFID 157 208 BY SIMILARITY.
CC FT DISULFID 263 311 BY SIMILARITY.
CC FT DISULFID 353 403 BY SIMILARITY.
CC FT DISULFID 447 496 BY SIMILARITY.
CC FT DISULFID 538 590 BY SIMILARITY.
CC FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 725 725 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 824 824 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 848 848 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 875 875 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1021 1021 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1029 1029 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1106 1106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASPLIC 1179 1182 Missing (in isoform 2).
FT FTId=VSP_002592.
SQ SEQUENCE 1259 AA; 140934 MW; 0F12A7C4415F3C08 CRC64;

Query Match 10.1%; Score 133; DB 1; Length 1259;
Best Local Similarity 24.4%; Pred. No. 0.024;
Matches 57; Conservative 31; Mismatches 96; Indels 50; Gaps 9;

OY 10 LLLVLTALLPAAQGNKVVYLGKGGDYELCTFASQKKSLQ--FHWKNSNOKITLNGQS 67
DB 507 ILANLVQVEATQITQPPSTIEKRGARTFTQASFPDSLQASITWKGDR----- 557
OY 68 FLTKGPSKLRADRSRLMDQGNFPIIKNLKEDSDTYIC---EVEDQKEVOLLVF 123
DB 558 -----DLQGRGSDKVFIEDG--LVYSLSVSDGQDYSCVASTELDEVESRAQLLV 608
OY 124 GLTN-----SDTHLLQGOSLPTLTSPSSSPSVQCRP-----RKNITGGKTL 169
DB 609 GSPGPVPHLELSDRHLLKQSQVHLW-----SPAEDHNSPIEKYDIEFEDKEMAPEKMF 662
OY 170 SVSGLLELDGSGTCTVQLQNGKVE--FKIDIVRASALPAPPGSALPDPOQA 221
DB 663 SLGRV---PQNOTSTLKLSPYVHYTRVAINKKGPEPSVSETVVPEAA 712

RESULT 17
CAML_MOUSE
ID CAML_MOUSE STANDARD; PRT; 1260 AA.
AC P1627;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule L1 precursor (N-CAM L1).
GN L1CAM OR CAML1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
[1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=88318924; PubMed=3412448;
RA Moos M., Tacke R., Scherer H., Teplow D., Fruh K., Schachner M.;
RT "Neural adhesion molecule L1 as a member of the immunoglobulin
RL superfamily with binding domains similar to fibronectin.";
RL Nature 334:701-703(1988).
CC -1- FUNCTION: Cell adhesion molecule with an important role in the
CC development of the nervous system. Involved in neuron-neuron
CC adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
CC to axonin on neurons.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC L1/neurofascin/NG2CAM family.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X12875; CA43368.1; -.
CC PIR; S05479; S05479.
CC HSSP; P20241; ICFB.
CC MGI; MGI:96721; L1cam.

```

```

DR GO: 0007411; P: axon guidance; IMP.
DR InterPro: IPR008957; FN_III-like.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003598; Ig_C2.
DR Pfam: PF00041; fn3; 4.
DR Pfam: PF00047; Ig; 6.
DR SMART: SM00060; FN3; 2.
DR SMART: SM00408; IGC2; 5.
DR PROSITE: PS50835; IG_Like; 6.
DR Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;
KM Transmembrane; Repeat; Immunoglobulin domain; Signal.
KT SIGNAL 1
FT CHAIN 20 1260 NEURAL CELL ADHESION MOLECULE L1.
FT DOMAIN 20 1123 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1124 1146 POTENTIAL.
FT DOMAIN 1147 1260 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 35 130 IG-LIKE C2-TYPE 1.
FT DOMAIN 138 225 IG-LIKE C2-TYPE 2.
FT DOMAIN 239 327 IG-LIKE C2-TYPE 3.
FT DOMAIN 332 419 IG-LIKE C2-TYPE 4.
FT DOMAIN 424 506 IG-LIKE C2-TYPE 5.
FT DOMAIN 517 600 IG-LIKE C2-TYPE 6.
FT DOMAIN 827 994 FIBRONECTIN TYPE-III 1.
FT DOMAIN 932 994 FIBRONECTIN TYPE-III 2.
FT DOMAIN 1032 1094 FIBRONECTIN TYPE-III 3.
FT SITE 553 555 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 564 564 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 57 113 BY SIMILARITY.
FT DISULFID 157 208 BY SIMILARITY.
FT DISULFID 263 311 BY SIMILARITY.
FT DISULFID 353 403 BY SIMILARITY.
FT DISULFID 447 496 BY SIMILARITY.
FT DISULFID 538 590 BY SIMILARITY.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 725 725 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 824 824 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 848 848 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 875 875 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 978 978 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1022 1022 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1073 1073 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1107 1107 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1260 AA; 140968 MW; 228B57001CB2A538 CRC64;

Query Match 9.9%; Score 131; DB 1; Length 1260;
Best Local Similarity 23.9%; Pred. No. 0.034;
Matches 56; Conservative 33; Mismatches 95; Indels 50; Gaps 9;

```

```

QY 170 SVSGLQLDSGTWCTVLOKQKVE--FKIDIVPRASALPAPPTGSALEDPQTA 221
DB 663 SLGKV---PGNQSTTLKLSPVHYTFRVTAINKYGPGEPSVSESVTPPEAA 712

RESULT 18
LAMP_HUMAN STANDARD; PRT; 338 AA.
ID LAMP_HUMAN
AC Q13449;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Limbic system-associated membrane protein precursor (LSAMP).
GN LSAMP OR LAMP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxId=9606;
RX MEDLINE=96235133; PubMed=8666243;
RA Pimenta A.F., Fischer I., Levitt P.;
RT "cDNA cloning and structural analysis of the human limbic-system-
associated membrane protein (LAMP).";
RL Gene 170:189-195(1996).

CC -1- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
CC MATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: Expressed on limbic neurons and fiber tracts
CC as well as in single layers of the superior colliculus, spinal
CC chord and cerebellum.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IgION
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U41901; AAC50569.1; -
CC PIR; JCA4776; JCA4776.
CC Genew; HGNC:6705; LSAMP.
CC MIM; 603241; -
DR GO: 0007399; P: neurogenesis; TNS.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003598; Ig_C2.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00408; IGC2; 2.
DR PROSITE: PS50835; IG_Like; 3.
KM Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KM Repeat; Signal; Lipoprotein.
FT CHAIN 1 28 POTENTIAL.
FT SIGNAL 29 315 LIMBIC SYSTEM-ASSOCIATED MEMBRANE
FT PROPEP 316 338 PROTEIN.
FT DOMAIN 29 122 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 132 214 IG-LIKE C2-TYPE 1.
FT DOMAIN 219 304 IG-LIKE C2-TYPE 2.
FT DISULFID 53 111 IG-LIKE C2-TYPE 3.
FT DISULFID 153 197 POTENTIAL.
FT DISULFID 239 290 POTENTIAL.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).

```


FT	CARBOHYD	300	300	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	315	315	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	LIPID	315	315	GPI-anchor amidated asparagine (Potential).
SQ	SEQUENCE	338 AA;	37308 MW;	03455F286DF5D92F CRC64;
Query Match		9.8%;	Score 129;	DB:1; Length 338;
Best Local Similarity		23.2%;	Pred. No. 0.0095;	
Matches		57;	Conservative 40;	Mismatches 91; Indels 58; Gaps 11
OY	10	LLVLVLQLLLPAA-----	10GKNVVLGGKGDPELCTASOKKSIFPHKNSNQIKI	61
DB	14	LVLLRLCLLPGLGVLVRSVDVDFRKGIDNTIVRGDTRILACVLDEKNS-KVAMLANSGILF		72
OY	62	LGNQGSFLTKGSPKLNDR--	DSRSLMDQGNFPLIKNLKIEDSDTYICEVEDQK--	116
DB	73	AGHD-----	KMSLDPRVLEKRHSL-----EYSLRIQGVADVDEGSYCSVOTQHEPKT	121
OY	117	-EVOLLVVG-----	LRANSPTHLLQGSLLTLTESPPGSSPSVQCR--SPRGKNIQGGT-	168
DB	122	SGVYLIIVGVPPKISNISSDVTYNESNVTLYCMANGREPEVITWHLLPTGREFGEERY		181
OY	169	LSVSGLELSDSGTWTCTVLQ-----	NQKVEEKIDIVP-----	203
DB	182	LEILGITHRQSGKECKAANEVSADVKOVKVTAVNPITTESKSNREATTGQASLCKBA		241
OY	204	SALPAP	209	
DB	242	SAVPAP	247	
RESULT 19				
ID	NCM2_MOUSE	STANDARD;	PRT;	837 AA.
AC	035136;	035962;		
DT	15-JUL-1998	(Rel. 36, Created)		
DT	15-JUL-1998	(Rel. 36, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Neural cell adhesion molecule 2 precursor (N-CAM 2) (RB-8 neural cell adhesion molecule) (R4812).			
DN	NCAM2 OR CCAM OR RNCAM.			
OS	Mus musculus (Mouse).			
OC	Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.			
OX	NCBI_TaxId=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).			
RC	STRAIN=BAJB/c; TISSUE=Olfactory neuroepithelium;			
RX	MEDLINE=97368238; PubMed=9221781;			
RA	Yoshihara Y., Kawasaki M., Tamada A., Fujita H., Hayaishi H.,			
RA	Kagamiyama H., Mori K.,			
RT	"OCAM: A new member of the neural cell adhesion molecule family related to zone-to-zone projection of olfactory and vomeronasal axons."			
RT	J. Neurosci. 17:5830-5842(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A. (ISOFORM SHORT).			
RP	STRAIN=C57BL/6J; TISSUE=Olfactory epithelium;			
RC	MEDLINE=97476194; PubMed=9334170;			
RX	Alentiu M., Bohm S.;			
RA	"Identification of a novel neural cell adhesion molecule-related gene with a potential role in selective axonal projection."			
RT	J. Biol. Chem. 272:26083-26086(1997).			
RL	[1]			
CC	-1- FUNCTION: May play important roles in selective fasciculation and zone-to-zone projection of the primary olfactory axons.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein (long isoform) and attached to the membrane by a GPI-anchor (short isoform).			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=Long;			
CC	IsoId=O35136-1; Sequence=Displayed;			
CC	Name=Short;			
CC	IsoId=O35136-2; Sequence=VSP_002590;			

```
CC -I- TISSUE SPECIFICITY: Expressed in subsets of both olfactory and vomeronasal neurons in a zone-specific manner.
CC -I- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -I- SIMILARITY: Contains 2 fibronectin type III domains.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements/ send an email to license@ebi.ac.uk).
CC -----
DR EMBL; AF001287; AAB69125.1; -
DR EMBL; AF001286; AAB69124.1; -
DR EMBL; AF016619; AAC53375.1; -.
DR MGI; MG1:97282; Ncam2.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003598; Ig_C2.
DR Pfam; PF00047; fn3; 2.
DR Pfam; PF00047; Ig; 5.
DR SMART; SMO0060; IG3; 2.
DR SMART; SMO0408; IGc2; 5.
DR PROSITE; PS50835; IG_LIKE; 5.
KW Cell adhesion; Transmembrane; Glycoprotein; Repeat; Immunoglobulin domain; Signal; GPI-anchor; Alternative splicing.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 837 NEURAL CELL ADHESION MOLECULE 2.
FT DOMAIN 20 697 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 698 718 POTENTIAL.
FT DOMAIN 719 837 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 21 108 IG-LIKE C2-TYPE 1.
FT DOMAIN 113 202 IG-LIKE C2-TYPE 2.
FT DOMAIN 208 297 IG-LIKE C2-TYPE 3.
FT DOMAIN 302 396 IG-LIKE C2-TYPE 4.
FT DOMAIN 401 491 IG-LIKE C2-TYPE 5.
FT DOMAIN 482 581 FIBRONECTIN TYPE-III 1.
FT DOMAIN 594 678 FIBRONECTIN TYPE-III 2.
FT DISULFID 42 93 PROBABLE.
FT DISULFID 136 186 PROBABLE.
FT DISULFID 232 281 PROBABLE.
FT DISULFID 322 380 PROBABLE.
FT DISULFID 422 475 PROBABLE.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 694 837 TLFGALGAIIGGVALLLIIVTVSGFFIQCGLLMC ITRMGCKSSGSSEKELEGKAAVYLDQSKDEIVMRT DERTIHEDSDPVNEBETTPLEPECLPLEENKGKVINA ETIRIKSYNDIIOGSKEDIKA -- NCCENAKNGGCSNH LNAVGTFVITMSLSCLF (In Isoform Short).
SQ SEQUENCE 837 AA; 93203 MW; 70473B053A2D65A5 CRC64; /ftid=vsp_002590.

Query Match 9.6%; Score 127; DB 1; Length 837;
Best Local Similarity 20.3%; Pred. No. 0.041;
Matches 48; Conservative 44; Mismatches 82; Indels 62; Gaps 10;
```

QY 117 E-VOLVLEGLTANSDFHLLQGOS-----LTLTLESPGSSPSVQC----- 155
 DB 294 AFIQVFA-----QPHILQKNETTSENGHVTLLYCEABEPVETIKRAIDGWFSSEG 346
 QY 156 -RSPRG-----KNIGGKTLVSQLELDOSGTWCTVLAQ-----NOKKVEFKIDIVP 202
 DB 347 DKSPDGEIEVKGQHGRLHIDRVKLSDSGRYDEASNRIGCHQRSHMLIEVAPK 402

RESULT 20
 LAMP RAT
 ID LAMP RAT STANDARD; PRT; 338 AA.
 AC 062813;
 DT 01-NOV-1997 (Rel. 35, 'Created')
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Limbic system-associated membrane protein precursor (LSAMP).
 GN LSAMP OR LAMP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-49.
 RC TISSUE=Hippocampus;
 RX MEDLINE=95374785; PubMed=7646886;
 RA Pimenta A.F., Zukareva V., Barde M.F., Retnoso B.S., Grimley C.,
 RA Henzel W., Fischer I., Levitt P.;
 RT "The limbic system-associated membrane protein is an Ig superfamily
 member that mediates selective neuronal growth and axon targeting";
 RL Neuron 15:287-297(1995).
 CC -1- FUNCTION: MEDIATES SELECTIVE NEURONAL GROWTH AND AXON TARGETING.
 CC CONTRIBUTES TO THE GUIDANCE OF DEVELOPING AXONS AND REMODELING OF
 CC NATURE CIRCUITS IN THE LIMBIC SYSTEM. ESSENTIAL FOR NORMAL GROWTH
 CC OF THE HYPOCAMPAL MOSSY FIBER PROJECTION.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: Expressed mostly by neurons comprising limbic-
 CC associated cortical and subcortical regions that function in
 CC cognition, emotion, memory, and learning.
 CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT E15-16, AT STAGE E20 IT IS
 CC DETECTED IN PRESUMPTIVE CORTEX, MEDIAL LIMBIC AREAS OF THE
 CC THALAMUS AND HYPOTHALAMUS. IN THE ADULT, IT IS FOUND IN
 CC HYPOTHALAMUS, PERIRHINAL CORTEX, AMYGALA AND MEDIAL THALMIC
 CC REGION.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IgION
 CC family.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcement/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; U31554; AAA86120.1; -
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_c2.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00408; IgC2; 2.
 DR PROSITE; PSS0835; IG-LIKE; 3.
 KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 KW Repeat; Signal; Lipoprotein.
 FT FT SIGNAL 1 28
 FT CHAIN 29 315
 FT PROPER 316 338 LIMBIC SYSTEM-ASSOCIATED MEMBRANE
 FT DOMAIN 29 122 PROTEIN.
 FT DOMAIN 132 214 REMOVED IN MATURE FORM (POTENTIAL).
 FT DOMAIN 219 304 IG-LIKE C2-TYPE 1.
 FT DISULFID 53 111 IG-LIKE C2-TYPE 2.
 FT DISULFID 153 197 POTENTIAL.
 FT POTENTIAL.

FT DISULFID 239 290 POTENTIAL.
 FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 279 279 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT LIPID 315 315 GPI-anchor amidated asparagine (Potential).
 SQ SEQUENCE 338 AA; 37324 MW; 0B76AFDD68A39B86 CRC64;

Query Match 9.6%; Score 126; DB 1; Length 338;
 Best Local Similarity 22.8%; Pred. No. 0.016;
 Matches 56; Conservative 41; Mismatches 91; Indels 58; Gaps 11;

QY 10 LLLVQLALLPAA-----TGANKVVLGKGDVLELTCTASQKSIQPHWNSNQIKI 61
 DB 14 LVILRLCLLPGLPVRSDVFNKGTNTVRGQDTALLCQVEDKIS-KVAMLNRSGLIF 72
 QY 62 LGNGSFLTKGSPKLNDRD-DSRSLMDQGNPLIINKLKIEDSTYICEVDQKE--- 116
 DB 73 AGHD-----KMSLDPRVLEKRRAL-----EYSLRIQKVVDYDEGSYCSYQTOHEPXT 121
 QY 117 -EYQLLVFG-----LTANSDFHLLQGOSLTLTLESPGSSPSVQC--SPRGKNIGGKT- 168
 DB 122 SQYLLLVQVPPKISNISSDVTNKGSNVTLVCMANGRPVITWRHLTPLGREFESEEY 181
 QY 169 LSVSQLELDOSGTWCTVLAQ-----NOKKVEFKIDIVP-----RA 203
 DB 182 LEILGTRGOSGKVECKANEVSADVKQKVTYVNPRTITSSKSNATTTGRASLKCA 241
 QY 204 SALPAP 209
 DB 242 SAVPAP 247

RESULT 21
 LAMP RAT
 ID LAMP RAT STANDARD; PRT; 739 AA.
 AC P29534;
 DT 01-APR-1993 (Rel. 25, 'Created')
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vascular cell adhesion protein 1 precursor (V-CAM 1).
 GN VCAM1 OR VCAM-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=92181437; PubMed=1371918;
 RA Heselin C., Moy P., Tizard R., Chisholm P., Williams C., Wysk M.,
 RA Barkly L., Miyake K., Kincade P., Lobb R.;
 RT "Cloning of murine and rat vascular cell adhesion molecule-1";
 RL Biochem. Biophys. Res. Commun. 183:163-169(1992).
 CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
 CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
 CC INTEGRIN VIA4 ON LEUKOCYTES. AND MEDIATES BOTH ADHESION AND SIGNAL
 CC TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A
 CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
 CC EMIGRATION TO SITES OF INFLAMMATION.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as
 CC well as on macrophage-like and dendritic cell types in both normal
 CC and inflamed tissue.
 CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1db-sib.ch/announce/> or send an email to license@1db-sib.ch).

EMBL; M84488; AAA42332.1; -

DR PIR; J50675; J50675.

DR HSSP; P19320; 1YCA.

DR InterPro; IPR003987; ICAM-VCAM-1.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR003989; VCAM-1.

DR Pfam; PF00047; Ig; 5.

DR PRINTS; PR01472; ICAMVCAM1.

DR PRINTS; PR01474; VCAM1.

DR SMART; SM00408; IGC2; 3.

DR PROSITE; PS50835; Ig_Like; 5.

KW Immunoglobulin domain; Glycoprotein; Cell adhesion; Transmembrane; Repeat; Signal.

FT SIGNAL 1 24 PROBABLE.

FT CHAIN 25 739 VASCULAR CELL ADHESION PROTEIN 1.

FT DOMAIN 25 698 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 699 720 POTENTIAL.

FT DOMAIN 721 739 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 25 111 IG-LIKE C2-TYPE 1.

FT DOMAIN 119 212 IG-LIKE C2-TYPE 2.

FT DOMAIN 223 309 IG-LIKE C2-TYPE 3.

FT DOMAIN 312 397 IG-LIKE C2-TYPE 4.

FT DOMAIN 408 506 IG-LIKE C2-TYPE 5.

FT DOMAIN 514 595 IG-LIKE C2-TYPE 6.

FT DOMAIN 601 682 IG-LIKE C2-TYPE 7.

FT DISULFID 47 95 BY SIMILARITY.

FT DISULFID 52 99 BY SIMILARITY.

FT DISULFID 137 195 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 650 650 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 650 650 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 739 AA; 81246 MW; 560855A1A1B100C CRC64;

Query Match 9.4%; Score 124; DB 1; Length 739;

Best Local Similarity 23.4%; Pred. No. 0.059; Indels 40; Gaps 5;

Matches 43; Conservative 30; Mismatches 71;

QY 33 KGDVETLTCTASQKSIQFMKNSNQIKLGNOSFLTKGSKINDRASHRSLMDQNF 92

DB 238 EGAIVTWCASEGLPAPRIFWSK-----LDNGVLQL-----SGNA 274

QY 93 PLIKNKLIKEDSDTYICE---VEDQKEVOLV-----FGLTANSPDTHLLQGOSLTLT 142

DB 275 TLTLIAMEDESGIYCEGVNLVGRDKTEVELIVQEKFTVDISPGSQAQVGDSDVLT 334

QY 143 LESPPGSSPSVQCSPRKNIQ-----DKTISVSGLEHLDGSGTCTVLQNKQKVEF 195

DB 335 CAAGCDSPPSWRTQDTSPLNGEVDEGATSTLTLSVGEVDEHSHCTVTCQRKLEK 394

QY 196 KIDI 199

DB 395 TIQV 398

RESULT 22

KILO RAT STANDARD; PRT; 348 AA.

AC Q9Z0J8;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Kilon protein precursor (kindred of Iglon).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

OX [1]

RN SEQUENCE FROM N.A., AND SEQUENCE OF 32-62.

RP MEDLINE=9915207; PubMed=10075727;

RX Funatsu N., Miyata S., Kumanogoh H., Shigeta M., Hamada K., Endo Y., Sakawa Y., Maekawa S.;

RA "Characterization of a novel rat brain glycosylphosphatidylinositol-anchored protein (Kilon), a member of the Iglon cell adhesion molecule family.";

RT J. Biol. Chem. 274:8224-8230(1999).

RL -1- FUNCTION: CELL-ADHESION (POTENTIAL).

CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.

CC -1- TISSUE SPECIFICITY: Highly expressed in brain.

CC -1- PTM: Glycosylated.

CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. Iglon family.

CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1db-sib.ch/announce/> or send an email to license@1db-sib.ch).

EMBL; AB017139; BAA75649.1; -

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_c2.

DR Pfam; PF00047; Ig; 3.

DR SMART; SM00408; IGC2; 2.

DR PROSITE; PS50835; Ig_Like; 3.

KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor; Repeat; Signal.

FT SIGNAL 1 31 KILO PROTEIN.

FT CHAIN 32 348 REMOVED IN MATURE FORM (POTENTIAL).

FT PROPEP 32 128 IG-LIKE C2-TYPE 1.

FT DOMAIN 133 215 IG-LIKE C2-TYPE 2.

FT DOMAIN 219 307 IG-LIKE C2-TYPE 3.

FT DISULFID 54 112 POTENTIAL.

FT DISULFID 154 197 POTENTIAL.

FT DISULFID 239 291 POTENTIAL.

FT CARBOHYD 67 67 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 348 AA; 37858 MW; 37E90D1C7D24ACAB CRC64;

Query Match 9.3%; Score 122; DB 1; Length 348;

Best Local Similarity 23.0%; Pred. No. 0.033;

Matches 53; Conservative 35; Mismatches 94; Indels 48; Gaps 10;

QY 20 PAATGKRVVLGKKDVELTCTASQKSIQFMKNSNQIKLGNOSFLTKG-----SK 75

DB 35 PMAAVDNLV--RKGDPAVRCYLEDGAS--KGAMLNRSIIIFAG--GDKWSVDPRAVIST 89

QY 76 LINDRADSRRLMDQGNFPLIKNKLIKEDSDTYICEVEDQKEVOLVFGTLA----- 127

DB 90 LNRK-----DYSIQIQNDVVDGDPYTCSTVOTHTPTMOVH-LTVQVPPRIYD 137

QY 128 -NSPDTHLLQGOSLTLTLESPPGSSPSVQCR--SPRGKNIQSGKTLVSQLELDGSGTWTG 184

DB 138 ISNMTINEGNTVTLTGLATGKPEPAISWHISBSAKFENGQYLDYGTTRQAGSYEC 197

QY 185 TV-----LQNKQKVEFKIDIVPRASALPAPT-----GSALPDP 218

DB 198 SAENDVFPDVKKVVRVNVNAPFTIQEIKSGTVPGRGLRCGAGVPP 247

Query Match	Best Local Similarity	Score	DB	Length
Matches 58; Conservative	35; Pident	No. 0.046; Mematches 100; Indels 36; Gaps 10;		

```

Oy 34 GDVTELECTRSQKKSIGIEFH-----KNSNOKIKINGQSFLLTKGPSKUNDBADSRSLAMDQ 89
Db 36 GGVELECSQVQADYDVLMKMKDRKQVDPPLPSTGSSLIINDSRALKALYDTRASS----- 91
Oy 90 GNPELLIKNLIKIEDSDTYICEV-----EDQKEVQLLVFG--LTRANSDHLLQGSGLTL 141
Db 92 -TYTLQIKDIQETDAGVYOCOVILIGNNKITAEVDLQVRPAPVISDSTSLVSEGOAV 150
Oy 142 TLSEPPSS--SPSVQGRS-----PRKNIQGKTLISVSOLELQDSQTMWCTYQLQN----- 169
Db 151 RLECYAGCYPRAPRSWNRENNALLPFGTSIVRGVGLXISIGKEDRRTYVC-VAENGVGK 209
Oy 190 --QKVEFKIDIVPRASALPAPPTGSAALP--DPQTASALPDDPPASAL 233
Db 210 GARNINAVEVEF-PPVITVPRPHLGQLQYDMDLCEHVEAYPPAIRWTL 257

RESULT 24.
CAML_FUGRU STANDARD; PRT; 1277 AA.
AC 098902:
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neutral cell adhesion molecule LI precursor (N-CAM LI) (LI-CAM).
GN LI.CAM.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
NCBI_TaxID=31033;
[1]
RN RA SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP TISSUE=Brain, and Muscle;
RC MEDLINE=98147998; PubMed=9479034;
RX Coutelle O., Nyakatura G., Tauden S., Elger G., Brenner S.,
RA Platzer M., Drescher B., Jouet M., Kennrich S., Rosenthal A.;
RT "The neural cell adhesion molecule LI: genomic organisation and
RT differential splicing is conserved between man and the pufferfish
RT Fugu."
RL Gene 208:7-15 (1998).
CC -!- FUNCTION: Cell adhesion molecule with an important role in the
CC development of the nervous system. Involved in neuron-neuron
CC adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
CC to axonin on neurons.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonym=Brain;
CC IsoId=Q98902-1; Sequence=displayed;
CC Name=2; Synonym=Muscle;
CC IsoId=Q98902-2; Sequence=VSP_050474, VSP_050475;
CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC LI/neurofascin/NgcAM family.
CC -!- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 5 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-eb.ch).
CC -----
DR EMBL; Z71926; CAA96469.1; -
DR EMBL; AF026198; AAC15580.1; -
DR PIR; T30532; T30532.
DR HSRP; P20241; ICBB.
DR InterPro; IPR008857; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; Ig-like.

```

```

DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00041; fn3; 5.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00060; FN3; 5.
DR SMART; SM00408; IGc2; 3.
DR PROSITE; PS50835; IG_LIKE; 6.
DR Cell adhesion; Glycoprotein; Transmembrane; Repeat;
KW Immunoglobulin domain; Signal; Alternative splicing.
FT SIGNAL 1 34
FT CHAIN 35 1277
FT DOMAIN 30 1135
FT TRANSMEM 1136 1156
FT DOMAIN 1157 1277
FT DOMAIN 51 140
FT DOMAIN 150 241
FT DOMAIN 256 344
FT DOMAIN 349 437
FT DOMAIN 443 528
FT DOMAIN 532 623
FT DOMAIN 628 726
FT DOMAIN 727 825
FT DOMAIN 826 932
FT DOMAIN 933 1028
FT DOMAIN 1029 1122
FT DISULFID 72 129
FT DISULFID 173 224
FT DISULFID 280 328
FT DISULFID 370 421
FT DISULFID 465 514
FT DISULFID 554 607
FT CARBOHYD 317 503
FT CARBOHYD 503 520
FT CARBOHYD 520 531
FT CARBOHYD 531 531
FT CARBOHYD 794 794
FT CARBOHYD 839 839
FT CARBOHYD 1035 1035
FT CARBOHYD 1046 1046
FT CARBOHYD 1068 1068
FT CARBOHYD 1083 1083
FT CARBOHYD 1108 1108
FT VASAPLIC 42 47
FT VASAPLIC 1190 1193
FT VASAPLIC 1277 AA; 141954 MW; 2845849BA9A42C27 CRC64;
SQ SEQUENCE 1277 AA; 141954 MW; 2845849BA9A42C27 CRC64;
Query Match 9.1%; Score 120; DB 1; Length 1277;
Best Local Similarity 24.7%; Pred. No. 0.23;
Matches 55; Conservative 34; Mismatches 92; Indels 42; Gaps 9;
QY 34 GDTVELTCTASQKSIQFHWKNSQIKILNQSGFLTKGPKLNDRADSRSLWDQGNFP 93
DB 363 GETVQLDQADGIGSPITW-----TVNGVPLSATSLPRSLTRESGS-- 405
QY 94 LIINKLKEDSDTYICEVEDOKER-----VQLVFGLTANSDFH-LLOGSLTL-- 141
DB 406 LILKDVIFGDTAIVQCSNKGITLANTNYVELPQQLITENGNTYTFEGGKALLEC 465
QY 142 -TLRSP-----PSSPVSQCRSPKGNKIQGKTLVSQLELQDSGTCTVLONOKVE 194
DB 466 ETFSPPKRYTWESSISLILADPRVNLITNG-GLEIANVSHDEGIYTCLVQGSNITSV 524
QY 195 FKIDIVPQASALPAPPTGASALPQTA---SALPDPAPASAL 233
DB 525 AEVEVLRNTVIL-SPPQALRLQPKTALFTCLVYTDPLSL 566
RESULT 25
CONT_CHICK STANDARD; PRT; 1010 AA.
AC P14781; P10450;
DT 01-MAR-1989 (Rel. 10, Created)

```

```

DT 01-APR-1990 (Rel. 14, Last sequence update)
DR 15-MAR-2004 (Rel. 43, Last annotation update)
DE Contractin precursor (Neural cell recognition molecule F11).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90180453; PubMed=2627374;
RA Brummenndorf T., Wolff J.M., Ratner F., Rathjen F.G.;
RT "Neural cell recognition molecule F11: homology with fibronectin type
RL Iii and immunoglobulin type C domains."
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=White leghorn;
RC MEDLINE=8908597; PubMed=3049624;
RA Ranscht B., Dours M.T.;
RT "Sequence of contractin, a 130-kD glycoprotein concentrated in areas
RL of interneuronal contact, defines a new member of the immunoglobulin
RN J. Cell Biol. 107:1561-1573(1988).
RN [3]
RP GPI-ANCHOR.
RX MEDLINE=89286606; PubMed=2735929;
RA Wolff J.M., Brummenndorf T., Rathjen F.G.;
RT "Neural cell recognition molecule F11: membrane interaction by
RL covalently attached phosphatidylinositol."
RN Biochem. Biophys. Res. Commun. 161:931-938(1989).
CC -1- FUNCTION: Mediates cell surface interactions during nervous system
CC development.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 4 fibronectin type Iii domains.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
CC framehifts in positions 1002 and 1007.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14877; CA33018.1; -.
DR EMBL; Y00813; CA68753.1; ALT_FRAME.
DR PIR; S01998; S01998.
DR HSSP; P40189; 1BQU.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00041; fn3; 4.
DR Pfam; PF00047; Ig; 6.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00408; IGc2; 5.
DR PROSITE; PS50835; IG_LIKE; 6.
KW Immunoglobulin domain; Membrane; Glycoprotein; Lipoprotein; Signal;
KW GPI-anchor; Cell adhesion; Repeat.
FT SIGNAL 1 19
FT CHAIN 20 984
FT PROPEP 985 1010
FT DOMAIN 133 123
FT DOMAIN 132 215
FT DOMAIN 232 317
FT DOMAIN 322 398
FT DOMAIN 404 491
FT DOMAIN 496 592
FT DOMAIN 593 599
FT DOMAIN 600 701
FT CONTRACTIN.
FT REMOVED IN NATURE FORM (POTENTIAL).
FT IG-LIKE C2-TYPE 1.
FT IG-LIKE C2-TYPE 2.
FT IG-LIKE C2-TYPE 3.
FT IG-LIKE C2-TYPE 4.
FT IG-LIKE C2-TYPE 5.
FT IG-LIKE C2-TYPE 6.
FT GLY/PRO-RICH.
FT FIBRONECTIN TYPE-III 1.

```

[illegible]

CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- TISSUE SPECIFICITY: Expressed in postmitotic neurons in 16-36 hour embryos, including those in the brain, cranial ganglia and olfactory placodes, and in all classes of spinal neurons.
CC	-1- DEVELOPMENTAL STAGE: Onset of expression correlates with the initiation of axonogenesis in 16-36 hour embryo.
CC	-1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC	LI/neurofascin/NgCAM family.
CC	-1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC	-1- SIMILARITY: Contains 5 fibronectin type III domains.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; X89204; CAA61490.1; --
DR	PIR; T30581; T30581.
DR	HSSP; P20241; ICFB.
DR	ZFIN; ZDB-GENE-980526-512; nad11.1.
DR	InterPro; IPRO08957; FN_III-like.
DR	InterPro; IPRO03961; FN_III.
DR	InterPro; IPRO03962; FNIII_subd.
DR	InterPro; IPRO07110; IG_II-like.
DR	InterPro; IPRO03598; IG_c2.
DR	Pfam; PF00041; fn3; 5.
DR	Pfam; PF00047; Ig; 6.
DR	PRINTS; PR00014; FNTYPEIII.
DR	SMART; SM00060; FN3; 5.
DR	SMART; SM00408; IGc2; 4.
DR	PROSITE; PS50835; IG_LIKE; 6.
KW	Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;
KM	Transmembrane; Repeat; Immunoglobulin domain.
FT	NON_TER 1
FT	DOMAIN <1 1054 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1055 1075 POTENTIAL.
FT	DOMAIN 1076 1197 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN <1 58 IG-LIKE C2-TYPE 1.
FT	DOMAIN 69 160 IG-LIKE C2-TYPE 2.
FT	DOMAIN 165 263 IG-LIKE C2-TYPE 3.
FT	DOMAIN 268 355 IG-LIKE C2-TYPE 4.
FT	DOMAIN 360 442 IG-LIKE C2-TYPE 5.
FT	DOMAIN 451 541 IG-LIKE C2-TYPE 6.
FT	DOMAIN 546 638 FIBONECTIN TYPE-III 1.
FT	DOMAIN 645 739 FIBONECTIN TYPE-III 2.
FT	DOMAIN 744 849 FIBONECTIN TYPE-III 3.
FT	DOMAIN 850 948 FIBONECTIN TYPE-III 4.
FT	DOMAIN 952 1029 FIBONECTIN TYPE-III 5.
FT	DISULFID 92 143 BY SIMILARITY.
FT	DISULFID 199 247 BY SIMILARITY.
FT	DISULFID 289 339 BY SIMILARITY.
FT	DISULFID 383 432 BY SIMILARITY.
FT	DISULFID 472 525 BY SIMILARITY.
FT	CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 421 421 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 959 959 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1002 1002 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 1027 1027 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 1197 AA; 132860 MW; 7CE1505EEBAC7B28 CRC64;
Query Match	9.0%; Score 118; DB 1; Length 1197;

Best Local Similarity 19.5%; Pred. No. 0.31;
Matches 44; Conservative 44; Mismatches 86; Indels 52; Gaps 8;

QY 9 HLLVLQALLPAATQGNKVVIGKGGDVELTCTASQKKSIOFHMKNSNOIKIINGNSF 68
D 259 HYTVTEA--PYWTRSPSEHLVAFGEVTRLDCKADGIPAPNTW-SINGPVSGTD--- 312
QY 69 LTKGPKLNDRADRSRLMDQGNFPLIKLKIEDSDTYICEVEDQKEVOLVFGJLAN 128
D 313 -----VDRRRV---SSGKLISNVEFSDTANYOCEAVNKGSIIL-----N 351
QY 129 SDTHLQGSJLTJ-----LESPPGSSPSVQCRS-----PRGNIQGK 167
D 352 THAVHVELPAQILTPDERLYQATAGQVMDLCRTFGSPFKIHWEILDISIPALSNKISQ 411
QY 168 T-----LSVSQLDQDSGWTCTVLCNOKKVEFKDIVRASAALPAP 209
D 412 TTNGSLKISNVEBDSNRYTCSVSETNKSISADVEVLRKIVGP 457

RESULT 27

OPCM_BOVIN STANDARD; PRT; 345 AA.
ID OPCM_BOVIN STANDARD; PRT; 345 AA.
AC P11834;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Opioid binding protein/cell adhesion molecule precursor (OBCAM)
DE (Opioid-binding cell adhesion molecule) (OPCM).
GN OPCML OR OBCAM OR OCAM.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=89251576; PubMed=2721489;
RA Schofield P.R., McFarland K.C., Hayflick J.S., Wilcox J.N.,
RA Cho T.M., Roy S., Lee N.M., Loh H.H., Seeburg P.H.;
RT "Molecular characterization of a new immunoglobulin superfamily
RT protein with potential roles in opioid binding and cell contact."
RL EMBO J. 8:489-495 (1989).
CC -1- FUNCTION: Binds opioids in the presence of acidic lipids; probably
CC involved in cell contact.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X12672; CA31192.1; -
DR PIR: S03199; S03199.
DR InterPro: IPR007110; IG-1like.
DR InterPro: IPR003598; IG_c2.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00408; IGc2; 2.
DR PROSITE: PS50835; IG_LIKE; 3.
KM Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal; Lipoprotein.
FT SIGNAL 1 27
FT CHAIN 28 322
FT PROPER 323 345
FT MOLECULE
FT REMOVED IN MATURE FORM (POTENTIAL).

FT DOMAIN 39 126 IG-LIKE C2-TYPE 1.
FT DOMAIN 126 219 IG-LIKE C2-TYPE 2.
FT DOMAIN 223 310 IG-LIKE C2-TYPE 3.
FT DISULFID 57 115 POTENTIAL.
FT DISULFID 157 202 POTENTIAL.
FT CARBOHYD 244 296 POTENTIAL.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 322 322 GPI-anchor amidated asparagine
FT SEQUENCE 345 AA; 37914 MW; DIECC8DE7DBC819 CRC64;

Query Match 8.9%; Score 117.5; DB 1; Length 345;
Best Local Similarity 22.4%; Pred. No. 0.073;
Matches 59; Conservative 43; Mismatches 100; Indels 61; Gaps 10;

QY 11 LVVLQALLPAATQGNKVVIG-----KKGDVLELTCTASQKKSIOFHMKNSN 57
D 14 LVVSLRLPLVPTGVPRSGDFTPRAMDVTYVQGESATLRCTIDRVT-RVAMLRNS 72
QY 58 QIKILGN-----QGSFLTKGPKLNDRADRSRLMDQGNFPLIKLKIEDSDTYICE 110
D 73 TILVAGNDRKSIDPRVILVNTPTQ-----YSIMIQNVVDYDGGPTTCS 116
QY 111 VE-----DQKEVOLVVG-----LTNSDTHLQGSJLTJLESPPGSSPSVQCRSPGKN 162
D 117 VQTNHFKTSRVHLIVQVPPQIMNISDVTVNEGSSVTLCLAIQGRDEFTVWHLISKE 176
QY 163 IQG-----GKLVSQLELDQSGTCTVLC-----NOKXVEFKDIVRASAALPAPRTGS 213
D 177 GQGVSEDEYLEISDIRDQSGEYECALNDVAPDVAKVKTIVNYPYIS--KAKTGV 234
QY 214 ALPPQTSALPDPASALPAA 236
D 235 SVGKGKILSC-----EASAVPMA 252

RESULT 28

LAMP_CHICK STANDARD; PRT; 338 AA.
ID LAMP_CHICK STANDARD; PRT; 338 AA.
AC Q98919;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Limbic system-associated membrane protein precursor (E19S) (CHLAMP,
DE G19-isoform).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97157768; PubMed=9004047;
RA Wilson D.J.A., Kim D.S., Clarke G.A., Marshall-Clarke S., Moss D.J.;
RT "A family of glycoproteins (GP55), which inhibit neurite outgrowth,
RT are members of the Ig superfamily and are related to OBCAM,
RT neurotrophin, LAMP and CEPU-1."
RL J. Cell Sci. 109:3129-3138 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97358596; PubMed=9215692;
RA Brumendorf T., Spaltmann F., Treubert U.;
RT "Cloning and characterization of a neural cell recognition molecule
RT on axons of the retinorecortical system and spinal cord."
RL Eur. J. Neurosci. 9:1105-1116 (1997).
CC -1- FUNCTION: Mediates selective neuronal growth and axon targeting.

Query Match	Best Local Match	Similarity	Score	DB 1	Length
8.9%	22.4%	Pred. NO. 0.078	93	Indels 56	Gaps 11
Matches 55	Conservative 41	Mismatches 93			
10	14	16	136	148	279
14	16	136	148	279	287
16	18	156	168	299	300
18	20	176	188	311	315
20	22	196	208	323	327
22	24	216	228	335	339
24	26	236	248	351	355
26	28	256	268	363	367
28	30	276	288	375	379
30	32	296	308	387	391
32	34	316	328	399	403
34	36	336	348	411	415
36	38	356	368	423	427
38	40	376	388	435	439
40	42	396	408	447	451
42	44	416	428	459	463
44	46	436	448	471	475
46	48	456	468	483	487
48	50	476	488	495	499
50	52	496	508	507	511
52	54	516	528	519	523
54	56	536	548	531	535
56	58	556	568	543	547
58	60	576	588	555	559
60	62	596	608	567	571
62	64	616	628	579	583
64	66	636	648	591	595
66	68	656	668	603	607
68	70	676	688	615	619
70	72	696	708	627	631
72	74	716	728	639	643
74	76	736	748	651	655
76	78	756	768	663	667
78	80	776	788	675	679
80	82	796	808	687	691
82	84	816	828	699	703
84	86	836	848	711	715
86	88	856	868	723	727
88	90	876	888	735	739
90	92	896	908	747	751
92	94	916	928	759	763
94	96	936	948	771	775
96	98	956	968	783	787
98	100	976	988	795	799
100	102	996	1008	807	811
102	104	1016	1028	819	823
104	106	1036	1048	831	835
106	108	1056	1068	843	847
108	110	1076	1088	855	859
110	112	1096	1108	867	871
112	114	1116	1128	879	883
114	116	1136	1148	891	895
116	118	1156	1168	903	907
118	120	1176	1188	915	919
120	122	1196	1208	927	931
122	124	1216	1228	939	943
124	126	1236	1248	951	955
126	128	1256	1268	963	967
128	130	1276	1288	975	979
130	132	1296	1308	987	991
132	134	1316	1328	999	1003
134	136	1336	1348	1011	1015
136	138	1356	1368	1023	1027
138	140	1376	1388	1035	1039
140	142	1396	1408</		

ID	VCAM1_HUMAN	STANDARD;	PTT;	739 AA.
AC	P19320;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Vascular cell adhesion protein 1 precursor (V-CAM 1) (CD106 antigen) (INCAM-100).			
CN	VCAM1 OR L1CAM.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiinda; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Umbilical vein;			
RX	MEDLINE=91016951; PubMed=1699207;			
RA	Polte T., Newman W., Gopal T.V.;			
RT	"Full length vascular cell adhesion molecule 1 (VCAM-1).";			
RT	Cytokine-induced endothelial protein that binds to lymphocytes.";			
RL	Cell 59:1203-1211(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91352090; PubMed=1715583;			
RX	Cybulsky M.I., Fries J.W.U., Williams A.J., Sulan P., Eddy R.,			
RA	Byers M., Shows T., Glimbert M.A. Jr., Collins T.;			
RT	"Gene structure, chromosomal location, and basis for alternative mRNA			
RT	splicing of the human VCAM1 gene.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:7859-7863(1991).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91201302; PubMed=1707873;			
RA	Hession C., Tizard R., Vassallo C., Schiffer S.B., Goff D., Moy P.,			
RA	Chi-Rosso G., Luhowskyj S., Lobb R., Osborn L.;			
RT	"Cloning of an alternate form of vascular cell adhesion molecule-1			
RT	(VCAM1).";			
RL	J. Biol. Chem. 266:6682-6685(1991).			
RN	[5]			
RP	SEQUENCE FROM N.A. AND VARIANTS PHE-318; ALA-384; ALA-413 AND			
RP	LEU-716.			
RA	Rider M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,			
RA	Rajkumar N.R., Toth E.J., Yi O., Nickerson D.A.;			
RL	Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A. (ISOFORM LONG).			
RC	TISSUE=Retinal pigment epithelium;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Bielow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Stapchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Diegleman M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaeetz T.E.,			
RA	Brownstein M.J., Uedin T.B., Toehiyuki S., Carninci P., Prange C.,			
RA	Bosak S.A., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,			
RA	Villalon D.K., Muzny D.M., Sodegryn E.J., Lu X., Gibbs R.A.,			
RA	Faney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzyzanski M.I., Skalka U., Smailus D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences.";			

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[7]

RA X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 25-226.

RA MEDLINE=95147978; PubMed=7531291;

RA Jones E.Y., Harlos K., Bottomley M.J., Robinson R.C., Driscoll P.C.,

RA Edwards R.M., Clements J.M., Dudgeon T.J., Stuart D.I.;

RT "Crystal structure of an integrin-binding fragment of vascular cell

RT adhesion molecule-1 at 1.8-A resolution.";

RL Nature 373:539-544(1995).

RA X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 25-220.

RA MEDLINE=95296382; PubMed=7539925;

RA Wang J.-H., Pepinsky R.B., Stehle T., Liu J.-H., Karpusas M.,

RA Osborn L.;

RT "Structure of a functional fragment of VCAM-1 refined at 1.9-A

RT resolution.";

RL Acta Crystallogr. D 52:369-379(1996).

CC -1- FUNCTION: IMPORANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION

CC IN LEUKOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1

CC INTEGRIN VIA ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL

CC TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A

CC PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE

CC EMIGRATION TO SITES OF INFLAMMATION.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Comment=Additional isoforms seem to exist;

CC Name=Long;

CC IsoId=P19320-1; Sequence=Displayed;

CC Name=Short;

CC IsoId=P19320-2; Sequence=VSP_002580;

CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as

CC well as on macrophage-like and dendritic cell types in both normal

CC and inflamed tissue.

CC -1- INDUCTION: By cytokines (e.g. IL-1, TNF-alpha).

CC -1- PTM: Sialoglycoprotein.

CC -1- DISEASE: May play an important role in the genesis of

CC atherosclerosis and rheumatoid arthritis.

CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.

CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD106 entry;

CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd106.htm".

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X53051; CAA37218.1; -;

DR EMBL; M30257; AAA51917.1; ALT_TERM.

DR EMBL; M73355; AAA61270.1; -;

DR EMBL; M60335; AAA61269.1; -;

DR EMBL; AF536818; AAM96190.1; -;

DR PIR; A41288; A41288.

DR PIR; B41288; B41288.

DR PDB; 1VCA; 15-SEP-95.

DR PDB; 1VSC; 20-JUN-96.

DR PDB; 1J39; 07-NOV-01.

DR Genew; HGNC:12663; VCAM1.

DR MIM; 192225; -;

DR InterPro; IPR003987; ICAM_VCAM-1.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig C2.

DR InterPro; IPR003989; VCAM-1.

DR Pfam; PF00047; Ig; 6.

DR PRINTS; PR01472; ICAMVCAM1.

DR PRINTS; PR01474; VCAM1.

DR SMART; SM00408; IGc2; 3.

DR PROSITE; PS00835; IG LIKE; 5.

KW Immunoglobulin domain, Glycoprotein, Cell adhesion, Transmembrane,

KW Repeat, Signal, Alternative splicing, Polymorphism, 3D-structure.

KW SIGNAL

FT 1 24

FT CHAIN

FT 25 739

FT DOMAIN

FT 25 698

FT TRANSMEM

FT 699 720

FT DOMAIN

FT 721 739

FT DOMAIN

FT 25 105

FT DOMAIN

FT 109 212

FT DOMAIN

FT 223 309

FT DOMAIN

FT 312 399

FT DOMAIN

FT 408 506

FT DOMAIN

FT 511 595

FT DOMAIN

FT 600 684

FT DISULFID

FT 47 95

FT DISULFID

FT 52 99

FT DISULFID

FT 137 195

FT CARBOHYD

FT 273 273

FT CARBOHYD

FT 365 365

FT CARBOHYD

FT 417 417

FT CARBOHYD

FT 463 463

FT CARBOHYD

FT 531 531

FT CARBOHYD

FT 561 561

FT VASPLIC

FT 310 402

FT VARIANT

FT 318 318

FT VARIANT

FT 384 384

FT VARIANT

FT 413 413

FT VARIANT

FT 716 716

FT STRAND

FT 26 30

FT STRAND

FT 34 38

FT TURN

FT 39 40

FT STRAND

FT 43 50

FT STRAND

FT 56 61

FT TURN

FT 62 63

FT STRAND

FT 70 74

FT TURN

FT 75 76

FT STRAND

FT 77 82

FT HELIX

FT 87 89

FT STRAND

FT 92 99

FT TURN

FT 100 101

FT STRAND

FT 102 114

FT STRAND

FT 120 123

Query Match 8.9%; Score 117; DB 1; Length 739;

Best Local Similarity 22.3%; Pred. No. 0.2;

Matches 47; Conservative 38; Mismatches 84; Indels 42; Gaps 7;

OY RHLLVLVQLALLPAAAT--QGKRVYLGKGGPVELTCTASQKSKIQFMKNSNQIKLIGNQ 65

DB 211 RQAVKELQVYISPKNTVYISVPSKTLQGGSVTWTCSEGLPAPEITW----- 258

OY 66 GSFLLTKPSKLNDRADSRSLMDQGNFPLIIKNLIKEDSPTYICE---VEDQKEEVQL 121

DB 259 -----SKKLDNGNLGHU---SGNATLTLILAMBMEDSGIYCEGVNLIGKRVKEVELI 307

OY 122 V----FGLTANSQTHLQ--GQSITLTLBSPGSSPSVQCRSPRGKNIQ-----GKT 168

DB 308 VQEKPFVTEISPGFRIAQIGDSVMLTCSYWGCESPFSFWRTOIDSPLSGKVRSEGTNST 367

OY 169 LSVSGLQDSDGCTWTCTVNLQNKVKVEFKIDI 199

Db 368 LTLSPVSENEHSLCTVTCGKKLEKGIQV 398

RESULT 30
OPCM_HUMAN STANDARD; PRT; 345 AA.

AC Q14982;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Opioid binding protein/cell adhesion molecule precursor (OBPCM)
DE Opioid-binding cell adhesion molecule (OBPCM).
GN OBPCM OR OBPCM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Occipital cortex;
RX MEDLINE=95237612; PubMed=7721093;
RA Sharg K.B., Lee N.M.;
RT "Cloning, sequencing and localization to chromosome 11 of a cDNA
RT encoding a human opioid-binding cell adhesion molecule (OBPCM).";
RL Gene 153:213-217(1995).
CC -1- FUNCTION: Binds opioids in the presence of acidic lipids; probably
CC involved in cell contact.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (by
CC similarity).
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IgIOLN
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L34774; AAA36387.1; -
DR PIR: J04025; J04025.
DR DR Genew; HGNC:8143; OPCML.
DR MIM: 600632; -
DR GO: GO:0005887; C: integral to plasma membrane; TAS.
DR GO: GO:0004985; F: opioid receptor activity; TAS.
DR GO: GO:0007155; P: cell adhesion; TAS.
DR GO: GO:0008038; P: neuronal cell recognition; TAS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_c2.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00408; IGC2; 2.
DR PROSITE: PS50835; IG_LIKE; 3.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal; Lipoprotein.
FT CHAIN 1 27
FT SIGNAL 1 27
FT CHAIN 28 322
FT PROPEP 323 345
FT DOMAIN 39 126
FT DOMAIN 136 219
FT DOMAIN 223 310
FT DISULFID 57 115
FT DISULFID 157 202
FT DISULFID 244 296
FT CARBOHYD 44 44
FT CARBOHYD 70 70
FT CARBOHYD 140 140
FT CARBOHYD 285 285
FT CARBOHYD 293 293
FT CARBOHYD 306 306

FT LIPID 322 322 GPI-anchor amidated asparagine
FT SEQUENCE 345 AA; 38007 MW; 87AD17BEA1AA3F4 CRC64;
SQ

Query Match 8.8%; Score 116.5; DB 1; Length 345;
Best Local Similarity 22.4%; Pred. No. 0.087;
Matches 59; Conservative 43; Mismatches 100; Indels 61; Gaps 10;

QY 11 LVLQALALPAATQGNKVVLG-----KKGDVELTCTASQKSKSIOFHKNKN 57
DB 14 LVVVSRLFLFVPTGVPVSGDAPPKAMDNVTVRQGSATRTCTIDRVT-RVAMLN 72
QY 58 QKILGN-----QGSFLTKGSKLNDRADSRSLMDQGNPLIKNKIEDSDTYICE 110
DB 73 TLIVAGNDKWSIDPRITILVNTPTQ-----YSIMQNDVVIDEGPYTS 116
QY 111 VE-----DQEEVQLVFG---LTNSDTHLLQGSITLTLSPSSPSVQCRSPRGN 162
DB 117 VQTDNHPKTSRVHLIYQVPPQIMNISDITVNEGSSVTLCLAGRPFTVTRHLSVKE 176
QY 163 IQG---GKTLVSQLELDQSGTWCTVYQ-----NQKYEFKIDVPPASALPAPPTGS 213
DB 177 GQGFVSEDEYLEISDIKRDQSGEYECALNDVAAPVRRKXTVNPYIS--KAKNTGV 234
QY 214 ALPDPQTASALPDPASALPAA 236
DB 235 SVGQKGLSC-----EASAVPWA 252

RESULT 31
LACH_DROME STANDARD; PRT; 359 AA.

AC Q24372;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lachesin precursor.
GN LAC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94038693; PubMed=8223276;
RA Karlstrom R.O., Walder L.P., Bastiani M.J.;
RT "Lachesin: an immunoglobulin superfamily protein whose expression
RT correlates with neurogenesis in grasshopper embryos.";
RL Development 118:509-522(1993).
CC -1- FUNCTION: MAY PLAY A ROLE IN EARLY NEURONAL DIFFERENTIATION AND
CC AXON OUTGROWTH.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (by
CC similarity).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ON DIFFERENTIATING NEURONAL CELLS
CC FROM THE ONSET OF NEUROGENESIS IN BOTH THE CENTRAL AND PERIPHERAL
CC NERVOUS SYSTEMS.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L13255; AAC37184.1; -
DR HSSP: P56276; ITLK.
DR FLYBASE: FBgn0010238; Lac.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_c2.

DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00408; IG2; 2.
 DR PROSITE; PS50835; IG LIKE; 3.
 KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
 Repeat; Signal; Lipoprotein.
 FT SIGNAL 1 25
 FT CHAIN 26 336
 FT PROPEP 337 359
 FT DOMAIN 29 130
 FT DOMAIN 135 221
 FT DOMAIN 226 317
 FT DISULFID 50 113
 FT DISULFID 157 204
 FT DISULFID 247 303
 FT CARBOHYD 92 140
 FT CARBOHYD 140 140
 FT LEU 336 336
 SQ SEQUENCE 359 AA; 39998 MW; D1F513E2B9D866E6 CRC64;
 Query Match 8.8%; Score 116.5; DB 1; Length 359;
 Best Local Similarity 22.3%; Pred. No. 0.091; 87; Indels 43; Gaps 7;
 Matches 47; Conservative 34; Mismatches 87; Indels 43; Gaps 7;
 QY 34 GDTVELTCTASQKSIQFHWKNSQIKILNQSGFLTKGPKLNDRADSRSLMDQGNFP 93
 DB 43 GGTVEFDCVQYAKEVNVLFKTDSDPVFLSTGSLVTKDSRFLRDPNPS-----TYK 97
 QY 94 LIINKLKEDSDTYICE-----VEDQKEVOLVFGLTANSDDLHLOGQSITLLESFPG 148
 DB 98 LQIDYIDETDAGTYCQVAVISTVHKVSAEKLVSRRPVID-----NSTGVVASEG 150
 QY 149 SSPSVQCS-----PRGNKIQGKTLVSQLELSDGFWTCVQLON- 189
 DB 151 SEVMECTASGYPPPTITWRRENNAILPTDSATTVGNTLRKSKYKEDRGTYIC-VADNG 209
 QY 190 -----QKVEFKIDIVPRASALPAPPTGSAL 215
 DB 210 VSKGDRNINVEVERAP-VITVPRRLGQAL 239
 RESULT 32
 PIGR RABIT STANDARD; PRT; 773 AA.
 AC P01832;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)
 DE [Contains: Secretory component].
 GN PIGR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84142246; PubMed=6322002;
 RA Mostov K.S., Friedlander M., Blobel G.;
 RT "The receptor for transepithelial transport of IgA and Igm contains
 RT multiple immunoglobulin-like domains."; Nature 308:37-43(1984).
 RL Nature 308:37-43(1984).
 RN [2]
 RP SEQUENCE OF 87-114 AND 410-428.
 RX MEDLINE=8828032; PubMed=313133;
 RA Frutiger S., Hughes G.J., Hanly W.C., Jaton J.-C.;
 RT "Rabbit secretory components of different allotypes vary in their
 RT carbohydrate content and their sites of N-linked glycosylation."; J. Biol. Chem. 263:8120-8125(1988)
 RL J. Biol. Chem. 263:8120-8125(1988)
 CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
 CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
 CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
 CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
 CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE

CC TRANSMEMBRANE SEGMENT.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Also secreted.
 CC -1- POLYMORPHISM: The sequence shown is that of allotype T62.
 CC -1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X00412; CAAG5118.1; -
 DR PIR; A02111; ORRBG
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR Pfam; PF00047; Ig; 5.
 DR PROSITE; PS50835; IG LIKE; 3.
 DR SMART; SM00409; IG; 5.
 KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
 KW Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 773
 FT CHAIN 19 615
 FT DOMAIN 19 647
 FT TRANSLEM 648 670
 FT DOMAIN 671 773
 FT DOMAIN 25 131
 FT DOMAIN 138 232
 FT DOMAIN 233 340
 FT DOMAIN 352 455
 FT DOMAIN 461 557
 FT DISULFID 46 115
 FT DISULFID 155 225
 FT DISULFID 260 324
 FT DISULFID 369 438
 FT DISULFID 478 538
 FT CARBOHYD 88 88
 FT CARBOHYD 108 108
 FT FT 418 418
 FT VARIANT 88 88
 FT VARIANT 94 94
 FT VARIANT 101 108
 FT VARIANT 110 110
 SQ SEQUENCE 773 AA; 83886 MW; DF2C44D2F1193C65 CRC64;
 Query Match 8.5%; Score 112.5; DB 1; Length 773;
 Best Local Similarity 23.4%; Pred. No. 0.47; 84; Indels 51; Gaps 7;
 Matches 52; Conservative 35; Mismatches 84; Indels 51; Gaps 7;
 QY 36 TVELTCTASQKSIQFHWKNSQIKILNQSGFLTKGPKLNDRADSRSLMDQGNFP 90
 DB 150 TWTTCFPTATRLKKS--FYKVEDELVLIISSSKAKDP-----RYKGRITLQIQS 202
 QY 91 ----NPLIINKLKIEDSDTYICE-----VEDQKEVOLVFGSL----- 125
 DB 203 TTAKFPTVTKHQLQNDAGQVYCCSGSDPRAEQQNDVRLTLPGLVGNLGSVTFECAL 262
 QY 126 ----TANSDDLHLOGQSITLLESPPGSSPPGKNIQGKTLVSQLELSDGFW 182
 DB 263 DSEDANAVASLDROGVGNVVIDSGCTIDPAFEGILFTKXENGFVSIVAGLRKEDTGY 322
 QY 183 TCTVQLON-----QKVEFKIDIVPRASALPAPPTGS 213
 DB 323 LCGVQSGSGSDGPTQLRQLFVNEIDVNSPPLTKKFPDGS 364
 RESULT 33
 CEAS HUMAN STANDARD; PRT; 702 AA.
 ID CEAS HUMAN
 AC P06731;

DT 01-JAN-1988 (Rel. 06, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43; Last annotation update)
DE Carcinoembryonic antigen-related cell adhesion molecule 5 precursor
DE (Carcinoembryonic antigen) (CEA) (Meconium antigen 100) (CD66e
DE antigen).
DE CEACAM5 OR CEA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90258861; PubMed=2342461;
RA Schrewe H., Thompson J., Bona M., Hefta L.J.F., Maruya A.,
RA Haasauer M., Shively J.E., von Kleier S., Zimmermann W.,
RT "Cloning of the complete gene for carcinoembryonic antigen: analysis
RT of its promoter indicates a region conveying cell type-specific
RT expression.";
RL Mol. Cell. Biol. 10:2738-2748(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86038876; PubMed=367012;
RA Beauchemin N., Benchimol S., Coutroyer D., Fuks A., Stanners C.P.,
RT "Isolation and characterization of full-length functional cDNA clones
RT for human carcinoembryonic antigen.";
RL Mol. Cell. Biol. 7:3221-3230(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89122014; PubMed=320478;
RA Barnette T., Goebel S.J., Nothurft M.A., Elting J.J.,
RT "Carcinoembryonic antigen family: characterization of cDNAs coding
RT for NCA and CEA and suggestion of nonrandom sequence variation in
RT their conserved loop-domains.";
RL Genomics 3:59-66(1988).
RN [4]
RP SEQUENCE OF 5-702 FROM N.A.
RX MEDLINE=87128144; PubMed=3814146;
RA Oikawa S., Nakazato H., Kosaki G.,
RT "Primary structure of human carcinoembryonic antigen (CEA) deduced
RT from cDNA sequence.";
RL Biochem. Biophys. Res. Commun. 142:511-518(1987).
RN [5]
RP SEQUENCE OF 331-702 FROM N.A.
RX MEDLINE=87204247; PubMed=3036371;
RA Zimmermann W., Ortlieb B., Friedrich R., von Kleier S.,
RT "Isolation and characterization of cDNA clones encoding the human
RT carcinoembryonic antigen reveal a highly conserved repeating
RT structure.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2960-2964(1987).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: Found in adenocarcinomas of endodermally
CC derived digestive system epithelium and fetal colon.
CC -1- PTM: COMPLEX IMMUNOREACTIVE GLYCOPROTEIN WITH A MW OF 180 kDa
CC COMPRISING 60% CARBOHYDRATE.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. CEA family.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like domains.
CC -1- DATABASE: NAME=CEA; NOTE=CD guide CD66e entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd66e.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M17303; AAB59513.1; -.
DR EMBL; M59262; AAB62835.1; ALT. SEQ.
DR EMBL; M59255; AAB62835.1; JOINED.
DR EMBL; M59256; AAB62835.1; JOINED.
DR EMBL; M59257; AAB62835.1; JOINED.

[illegible]

Db 544 --SPRLQSLNGNRLLTLLFNTRNDARAYVCCIONSVSANRSDPYTLVDLYGPDPIIISP 601
Qy 130 DTHLLOGSLTLTLESPPGSSPSVOCRRPRKNIQGGKTLVSQLELDGSGTWCTVLQN 189
Db 602 DSSYLSGANLNLWLSCHSASNPSPQYSWR1-NCIPQHTVLTALITPNNNGTACFV-- 657
Qy 190 QKKVEFKIDIVPAPSPAPPTGASALPDPPTASALPDPPASALPALALAVSFLGLGLG 249
Db 658 -----SNL-ATGRNNSIVKSIIVSASGIRSGLS-----AGATVGMIGVLVG 698
Qy 250 VACV 253
Db 699 VALI 702
RESULT 34
LAR DROME STANDARD, PRT: 2029 AA.
ID LAR DROME STANDARD, PRT: 2029 AA.
AC P16621;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein-tyrosine phosphatase Lar precursor (EC 3.1.3.48) (Protein-tyrosine-phosphate phosphohydrolase) (DLAR).
GN LAR.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90046860; PubMed=2554325;
RA Streuli M., Krueger N.X., Tsai A.Y.M., Saito H.;
RT "A family of receptor-linked protein tyrosine phosphatases in humans and Drosophila."
RL Proc. Natl. Acad. Sci. U.S.A. 86:8698-8702(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RA MEDLINE=96178473; PubMed=8598047;
RX Krueger N.X., van Vactor D., Wan H.I., Gelbart W.M., Goodman C.S., Saito H.;
RT "The transmembrane tyrosine phosphatase DLAR controls motor axon guidance in Drosophila."
RL Cell 84:611-622(1996).
CC -1- FUNCTION: It is possible that DLAR is a cell adhesion receptor. It possesses an intrinsic protein tyrosine phosphatase activity (PTPase). It controls motor axon guidance.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Selectively expressed in a subset of axons and pioneer neurons in the embryo.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 9 fibronectin type III domains.
CC -1- SIMILARITY: Contains 2 protein-tyrosine phosphatase domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
CC or send an email to license@ebi.ac.uk.
DR EMBL: M27700; AAA28668.1; -
DR EMBL: U36857; AAC47002.1; -
DR EMBL: U36849; AAC47002.1; JOINED.
DR EMBL: U36850; AAC47002.1; JOINED.
DR EMBL: U36851; AAC47002.1; JOINED.
DR EMBL: U36852; AAC47002.1; JOINED.
DR EMBL: U36853; AAC47002.1; JOINED.

DR EMBL: U36854; AAC47002.1; JOINED.
DR EMBL: U36855; AAC47002.1; JOINED.
DR EMBL: U36856; AAC47002.1; JOINED.
DR PIR: A36182; TDFELK.
DR HSSP: P28827; IRPM.
DR FLYBase; FBgn0000464; Lar.
DR GO: GO:0004725; P:protein tyrosine phosphatase activity; IDA.
DR GO: GO:0008045; P:motor axon guidance; IMP.
DR GO: GO:0006470; P:protein amino acid dephosphorylation; IDA.
DR InterPro: IPR008957; FN-III-like.
DR InterPro: IPR003961; FN-III.
DR InterPro: IPR003962; FNIII subd.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig C2.
DR InterPro: IPR000387; TYR_phosphatase.
DR InterPro: IPR00242; TYR_PP.
DR Pfam: PF00047; Ig; 3.
DR Pfam: PF00041; fn3; 9.
DR Pfam: PF00102; Y_phosphatase; 2.
DR PRINTS: PR00014; FNTYPEIII.
DR PRINTS: PR00700; PRTYPHPTASE.
DR SMART: SM00660; FN3; 9.
DR SMART: SM00408; IGC2; 3.
DR SMART: SM00194; PTPC; 2.
DR PROSITE: PS50835; IG_LIKE; 3.
DR PROSITE: PS50383; TYR_PHOSPHATASE_1; 2.
DR PROSITE: PS50055; TYR_PHOSPHATASE_2; 2.
DR PROSITE: PS50055; TYR_PHOSPHATASE_2; 2.
KW Hydroxylase; Receptor; Glycoprotein; Signal; Transmembrane; Cell adhesion; Immunoglobulin domain; Repeat.
FT SIGNAL 1 32
FT CHAIN 33 2029
FT DOMAIN 33 1377
FT TRANSMEM 1378 1402
FT TRANSMEM 1403 2029
FT DOMAIN 36 128
FT DOMAIN 140 224
FT DOMAIN 234 316
FT DOMAIN 320 417
FT DOMAIN 418 512
FT DOMAIN 513 607
FT DOMAIN 608 706
FT DOMAIN 707 809
FT DOMAIN 810 906
FT DOMAIN 907 1007
FT DOMAIN 1008 1102
FT DOMAIN 1103 1207
FT DOMAIN 1492 1738
FT DOMAIN 1781 2029
FT ACT_SITE 1670 1670
FT ACT_SITE 1961 1961
FT DISULFID 57 111
FT DISULFID 161 209
FT DISULFID 256 301
FT CARBOHYD 176 176
FT CARBOHYD 253 253
FT CARBOHYD 298 298
FT CARBOHYD 553 553
FT CARBOHYD 616 616
FT CARBOHYD 666 666
FT CARBOHYD 721 721
FT CARBOHYD 774 774
FT CARBOHYD 915 915
FT CARBOHYD 962 962
FT CARBOHYD 1183 1183
FT CARBOHYD 1304 1304
SO SEQUENCE 2029 AA; 229027 MM; 536AOC794D3DC800 CRC64;
Query Match 8.4%; Score 110.5; DB 1; Length 2029;
Best Local Similarity 23.0%; Pred. No. 2.2;
Matches 50; Conservative 37; Mismatches 75; Indels 55; Gaps 11;

QY 20 PAATGKVKVVGKKDPTVELTCTASQKKSIOFMKMSNOIKILGNQGSFLTKGPKSLNDR 79
 DB 140 PVTGGTGVTRVIEGVHVTMTCKAIGNPTNPIYV-INKQTKV-----DW 182
 QY 80 ADNRSRSLMDQGNPLIKNLIKEDSPYICEVEDQKEVQLVFGGLTANSDTHL-LOGGS 138
 DB 183 SNRYSLSKD--GF-LOJENREDQGRYECVAENS-----MGTEHSKATNLYKVR 231
 QY 139 LVTLESPP-----GSSPSVQCRS-----PRGXNIGGKTLT-----VSOL 174
 DB 232 VPTFSPRPETISEVMIGSNLNLSCIAVGSMPHVKKMKSGSDLTPEENEMPRGNVQLI 291
 QY 175 ELQDSGTCTCTVQGNQKVEFKIDIVPRASALPAPT 211
 DB 292 NIGESANYTCTIAASTGLQID-SVSVV-KVQSLPTAPT 326

RESULT 35
 VCAL_MOUSE
 ID VCAL_MOUSE STANDARD; PRT; 739 AA.
 AC P29533;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vascular cell adhesion protein 1 precursor (V-CAM 1).
 GN VCAM1 OR VCAM-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=FVB; TISSUE=Lung;
 MDLINE=92161437; PubMed=1371918;
 RA Herskron C., Moy P., Tizard R., Chisholm P., Williams C., Wysk M.,
 RA Berkley L., Miyake K., Kincaid P., Lobb R.;
 RT "Cloning of murine and rat vascular cell adhesion molecule-1.";
 RL Biochem. Biophys. Res. Commun. 183:163-169(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Lymph node;
 RX MEDLINE=93246254; PubMed=7683304;
 RA Araki M., Araki K., Vassalli P.;
 RT "Cloning and sequencing of mouse VCAM-1 cDNA.";
 RL Gene 126:261-264(1993).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=129; TISSUE=Embryo;
 MDLINE=94117008; PubMed=7507076;
 RA Cybulsky M.I., Allan-Motamed M., Collins T.;
 RT "Structure of the murine VCAM1 gene.";
 RL Genomics 18:387-391(1993).
 RN [4]
 RP SEQUENCE OF 1-693 FROM N.A. (ISOFORM 1).
 RC STRAIN=NIH Swiss, and 129/Sv;
 RA Kumar A.G., Dai Y.X., Kozak C.A., Mims M.P., Gotto A.M. Jr.,
 RA Ballantyne C.M.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=FVB; TISSUE=Lung;
 MDLINE=93232042; PubMed=7682556;
 RA Moy P., Lobb R., Tizard R., Olson D., Hession C.;
 RT "Cloning of an inflammation-specific phosphatidyl inositol-linked
 form of murine vascular cell adhesion molecule-1.";
 RL J. Biol. Chem. 268:8835-8841(1993).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=C57BL/6; TISSUE=Liver;
 RX MEDLINE=95015899; PubMed=7523515;
 RA Kumar A.G., Dai X.Y., Kozak C.A., Mims M.P., Gotto A.M.,
 RA Ballantyne C.M.;

RT "Murine VCAM-1. Molecular cloning, mapping, and analysis of a
 truncated form.";
 RT J. Immunol. 153:4088-4096(1994).
 RN [7]
 RP SEQUENCE OF 311-345 FROM N.A. (ISOFORM 2).
 RC STRAIN=FVB/N; TISSUE=Kidney;
 RX MEDLINE=93317595; PubMed=7687058;
 RA Terry R.W., Kwee L., Levine J.F., Labow M.A.;
 RT "Cytokine induction of an alternatively spliced murine vascular cell
 adhesion molecule (VCAM) mRNA encoding a
 glycosylphosphatidylinositol-anchored VCAM protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:5919-5923(1993).
 RN [8]
 RP SEQUENCE OF 1-21 FROM N.A.
 RC TISSUE=Endothelial cells;
 RA Korenaga R., Ando J., Tsuboi H., Kamiya A.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: IMPORTANT IN CELL-CELL RECOGNITION. APPEARS TO FUNCTION
 IN LEUCOCYTE-ENDOTHELIAL CELL ADHESION. INTERACTS WITH THE BETA-1
 INTEGRIN VLA4 ON LEUKOCYTES, AND MEDIATES BOTH ADHESION AND SIGNAL
 TRANSDUCTION. THE VCAM1/VLA4 INTERACTION MAY PLAY A
 PATHOPHYSIOLOGIC ROLE BOTH IN IMMUNE RESPONSES AND IN LEUKOCYTE
 EMIGRATION TO SITES OF INFLAMMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG ISOFORM) OR
 ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR (SHORT ISOFORM).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=Long;
 CC IsoId=P29533-1; Sequence=Displayed;
 CC Name=2; Synonyms=Short;
 CC IsoId=P29533-2; Sequence=VSP_002581, VSP_002582;
 CC -1- TISSUE SPECIFICITY: Expressed on inflamed vascular endothelium, as
 well as on macrophage-like and dendritic cell types in both normal
 and inflamed tissue.
 CC -1- PTM: The GPI-anchor is located on position 319 of isoform 2.
 CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M64487; AAA40545.1; -
 CC EMBL; X67783; CAA47989.1; -
 CC EMBL; L22355; AAA16921.1; -
 CC EMBL; L22301; AAA16921.1; JOINED.
 CC EMBL; L22349; AAA16921.1; JOINED.
 CC EMBL; L22350; AAA16921.1; JOINED.
 CC EMBL; L22351; AAA16921.1; JOINED.
 CC EMBL; L22352; AAA16921.1; JOINED.
 CC EMBL; L22353; AAA16921.1; JOINED.
 CC EMBL; L22354; AAA16921.1; JOINED.
 CC EMBL; L22350; AAA16920.1; -
 CC EMBL; L22301; AAA16920.1; JOINED.
 CC EMBL; L22349; AAA16920.1; JOINED.
 CC EMBL; U12878; AAB60659.1; ALT_SEQ.
 CC EMBL; U12879; AAB60660.1; ALT_SEQ.
 CC EMBL; U12880; AAB60661.1; ALT_SEQ.
 CC EMBL; U12874; AAB60662.1; ALT_SEQ.
 CC EMBL; U12875; AAB60663.1; ALT_SEQ.
 CC EMBL; U12883; AAB60664.1; ALT_SEQ.
 CC EMBL; U12881; AAA80010.1; ALT_SEQ.
 CC EMBL; U12882; AAA80011.1; ALT_SEQ.
 CC EMBL; U12875; AAA80012.1; ALT_SEQ.
 CC EMBL; U12872; AAA80013.1; ALT_SEQ.
 CC EMBL; U12876; AAA80014.1; ALT_SEQ.
 CC EMBL; U12873; AAA80015.1; ALT_SEQ.
 CC EMBL; U12877; AAA80016.1; ALT_SEQ.
 CC EMBL; U108431; AAA40546.1; -
 CC EMBL; U12884; AAA64832.1; -

[illegible]

DT 16-Oct-2001 (Rel. 40, Last sequence update)
 DE 10-Oct-2003 (Rel. 42, Last annotation update)
 DT Amalgam protein precursor.
 GN AMA OR BG:DS00276.6 OR CG2198.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=689028670; PubMed=3141062;
 RA Seeger M.A., Halfley L., Kaufman T.C.;
 RT "Characterization of amalgam: a member of the immunoglobulin
 RT superfamily from Drosophila.";
 RL Cell 55:569-600(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RA Celniker S.E., Pfeiffer B.D., Knafels J., Martin C.H., Mayeda C.A.,
 RA Palazzolo M.J.;
 RT "Complete sequence of the Antennapedia complex of Drosophila.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBD databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Gutter R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Broctier P.,
 RA Butts K.C., Buam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo S., Delcher A., Deng Z., Dugas A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dwyer-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liang X., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mohreri A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein D.R., Paclé J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yah R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin C.M., Venter J.C.;
 RL "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Chame M.,
 RA George R.A., Garin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;

```

RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (Potential).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23561; AAA28367.1; -.
DR EMBL: AE001572; AAD19797.1; -.
DR EMBL: AE003674; AAF54084.1; -.
DR EMBL: AY051911; AAK93335.1; -.
DR PIR: A11923; A11923.
DR FlyBase; FBgn0000071; Ama.
DR GO: GO:0005886; C:plasma membrane; IDA.
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00047; IG_3.
DR SMART; SM00408; IGC2; 2.
DR PROSITE; PSS00835; IG_LIKE; 3.
KW Immunoglobulin domain; Glycoprotein; Membrane; GPI-anchor; Signal;
KW Repeat.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 7 AMALGAM PROTEIN.
FT PROPEP ? 333 REMOVED IN MATURE FORM (POTENTIAL).
FT DOMAIN 25 128 IG-LIKE V-TYPE.
FT DOMAIN 139 223 IG-LIKE C2-TYPE 1.
FT DOMAIN 230 323 IG-LIKE C2-TYPE 2.
FT DISULFID 46 117 PROBABLE.
FT DISULFID 161 208 PROBABLE.
FT DISULFID 251 307 PROBABLE.
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 83 83 O -> K (IN REF. 1).
SQ SEQUENCE 333 AA; 36387 MW; F644753DE3DB25F1 CRC64;

Query Match 8.3%; Score 109.5; DB 1; Length 333;
Best Local Similarity 20.8%; Pred. No. 0.28;
Matches 49; Conservative 40; Mismatches 88; Indels 59; Gaps 10;

OY 10 LLLVLLQAL-----LPATQGNKVVLGKGDVTELTCTASQKSIQFHW-----KNSNQ 58
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 10 LIFPLAISLSVLSAPVISOISKOVVASVGSVEFNCTVEVQGLSVMAKRSESDTNS 69
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 59 IKI-----LGNO--GSFLTKGSPKLNDRADSRSLMDQGNFPLIKNKIEDSDTYIC 109
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 70 VVLISMNRLILPDQRVNVVTEGPKT-----GSAIFYFRIONIEVSMGPHYEC 117
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 110 EV-----EDQKEVOLL---VFGLTANSDTHLLOGOSILTLTLESPGSSPSV----- 153
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 118 QVLVSATEKYTKSLQIKTPPVIAENTPKSTLVEGONELTCHANGFKPKPTISWAREH 177
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
OY 154 QCRSPRKNIQGCKTSLVSQLELDQSGTWTCTVLQN-----OKYVEFKIDI 199
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB 178 NAWPAPGCHLAEFTLRIRSVHRMDRGYYC-IAQNEGQPRDKRLINVEVEFRQI 232
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

RESULT 37
ID CAML HUMAN STANDARD; PRT; 1257 AA.
AC P32056; O8TA87;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule L1 precursor (N-CAM L1) (CD171 antigen).

```

```

GN LICAM OR CAML1 OR MICS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92031698; PubMed=1932117;
RA Kobayashi M., Miura M., Asou H., Uyemura K.;
RT "Molecular cloning of cell adhesion molecule L1 from human nervous
RT tissue: a comparison of the primary sequences of L1 molecules of
RT different origin."
RL Biochim. Biophys. Acta 1090:238-240 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92120663; PubMed=1796655;
RA Hlavín M.L., Lemmon V.;
RT "Molecular structure and functional testing of human LICAM: an
RT interspecies comparison."
RL Genomics 11:416-423 (1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92329299; PubMed=1627459;
RA Reid R.A., Hemperly J.J.;
RT "Variants of human L1 cell adhesion molecule arise through alternate
RT splicing of RNA."
RL J. Mol. Neurosci. 3:127-135 (1992).
RN [4]
RP SEQUENCE FROM N.A.
RA Rosenthal A., Couelle O., Drescher B.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=97432815; PubMed=9286695;
RA Brenner V., Nyakatura G., Rosenthal A., Platzner M.;
RT "Genomic organization of two novel genes on human Xq28: compact head
RT to head arrangement of IDH gamma and TRAP delta is conserved in rat
RT and mouse."
RL Genomics 44:8-14 (1997).
RN [6]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98147998; PubMed=9479034;
RA Couelle O., Nyakatura G., Taudien S., Elgar G., Brenner S.,
RA Platzner M., Drescher B., Joutet M., Kenzwick S., Rosenthal A.;
RT "The neural cell adhesion molecule L1: genomic organization and
RT differential splicing is conserved between man and the pufferfish
RT Fugu."
RL Gene 208:7-15 (1998).
RN [7]
RP SEQUENCE OF 20-36.
RX MEDLINE=88298876; PubMed=3136168;
RA Wolf J.M., Frank R., Mujoo K., Spiro R.C., Reifeld R.A.,
RA Rathjen F.G.;
RT "A human brain glycoprotein related to the mouse cell adhesion
RT molecule L1."
RL J. Biol. Chem. 263:11943-11947 (1988).
RN [8]
RP SEQUENCE OF 332-371 FROM N.A.
RX MEDLINE=90353957; PubMed=2387585;
RA Djabali M., Matedel M.-G., Nguyen C., Roux D., Demengeot J.,
RA Denisot F., Moos M., Schachner M., Goridis C., Jordan B.R.;
RT "The gene encoding L1, a neural adhesion molecule of the
RT immunoglobulin family, is located on the X chromosome in mouse and
RT man."
RL Genomics 7:587-593 (1990).
RN [9]
RP SEQUENCE OF 353-1176 FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=92020233; PubMed=1923824;
RA Rosenthal A., Mackinnon R.N., Jones D.S.C.;
RT "PCR walking from microdissection clone M54 identifies three exons
RT from the human gene for the neural cell adhesion molecule L1

```

RT (CAM-L1).";
 RL Nucleic Acids Res. 19:5395-5401(1991).
 RN [10]
 RP SEQUENCE OF 809-1257 FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Fetsch R.D., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,
 Stajich M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ugin T.B., Toshitsuki S., Carinici P., Prange C.,
 Bata S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Small D.E.,
 Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [11]
 RP SEQUENCE OF 1030-1257 FROM N.A.
 RX MEDLINE=9113183; PubMed=1993895;
 RA Harper J.R., Prince J.T., Healy P.A., Stuart J.K., Nauman S.J.,
 Stalcup W.B.;
 RT "Isolation and sequence of partial cDNA clones of human L1: homology
 of human and rodent L1 in the cytoplasmic region.";
 RL J. Neurochem. 56:797-804(1991).
 RN [12]
 RP VARIANT HSAS TYR-264.
 RX MEDLINE=94004956; PubMed=8401576;
 RA Jouet M., Rozenthal A., Macfarlane J., Kenrick S., Donnai D.;
 RT "A missense mutation confirms the L1 defect in X-linked hydrocephalus
 (HSAS).";
 RL Nat. Genet. 4:331-331(1993).
 RN [13]
 RP VARIANT HSAS/MASA LEU-1194.
 RX MEDLINE=95187172; PubMed=7881431;
 RA Franssen E., Schander-Stumpel C., Vits L., Coucke P., van Camp G.,
 Willems P.J.;
 RT "X-linked hydrocephalus and MASA syndrome present in one family are
 due to a single missense mutation in exon 28 of the L1CAM gene.";
 RL Hum. Mol. Genet. 3:2255-2256(1994).
 RN [14]
 RP VARIANTS HSAS GLN-184 AND ARG-452, AND VARIANT MASA GLN-210.
 RX MEDLINE=95004608; PubMed=7920659;
 RA Jouet M., Rozenthal A., Armstrong G., Macfarlane J., Stevenson R.,
 Paterson J., Metzenberg A., Ionescu V., Temple K., Kenrick S.;
 RT "X-linked spastic paraplegia (SPG1), MASA syndrome and X-linked
 hydrocephalus result from mutations in the L1 gene.";
 RL Nat. Genet. 7:402-407(1994).
 RN [15]
 RP VARIANTS MASA GLN-210 AND ASN-598.
 RX MEDLINE=95004609; PubMed=7920660;
 RA Vits L., van Camp G., Coucke P., Franssen E., de Boule K.,
 Reijnders E., Korn B., Poucka A., Wilson G., Schander-Stumpel C.,
 Winter R.M., Schwartz C., Willems P.J.;
 RT "MASA syndrome is due to mutations in the neural cell adhesion gene
 L1CAM.";
 RL Nat. Genet. 7:408-413(1994).
 RN [16]
 RP VARIANTS HSAS/MASA SER-9; SER-121; LYS-309; PHE-768; LEU-941 AND
 RX MEDLINE=95282776; PubMed=7762552;
 RA Jouet M., Moncla A., Paterson J., McKewen C., Fryer A., Carpenter N.,
 Holmberg E., Wadell C., Kenrick S.;
 RT "New domains of neural cell-adhesion molecule L1 implicated in

RT X-linked hydrocephalus and MASA syndrome.";
 RL Am. J. Hum. Genet. 56:1304-1314(1995).
 RN [17]
 RP VARIANTS HSAS/MASA GLN-184; GLN-210; TYR-264; ARG-452; ASN-598 AND
 RX LEU-1194.
 RX MEDLINE=96151146; PubMed=8556302;
 RA Franssen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;
 RT "CRASH syndrome: clinical spectrum of corpus callosum hypoplasia,
 retardation, adducted thumbs, spastic paraparesis and hydrocephalus
 due to mutations in one single gene, L1.";
 RL Eur. J. Hum. Genet. 3:273-284(1995).
 RN [18]
 RP ERRTM.
 RA Franssen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;
 RL Eur. J. Hum. Genet. 4:126-126(1996).
 RN [19]
 RP VARIANTS HSAS/MASA/SPG1 SER-179 AND ARG-370.
 RX MEDLINE=96057511; PubMed=7562969;
 RA Ruiz J.C., Cuppens H., Legius E., Fyfe J.-P., Glover T., Marynen P.,
 Casman J.-J.;
 RT "Mutations in L1-CAM in two families with X linked complicated
 spastic paraplegia, MASA syndrome, and HSAS.";
 RL J. Med. Genet. 32:549-552(1995).
 RN [20]
 RP VARIANTS HSAS CYS-194 AND LEU-240.
 RX MEDLINE=97083379; PubMed=8929944;
 RA Gu S.-M., Orth U., Veske A., Enders H., Kluender K., Schloesser M.,
 Engel W., Schwinger E., Gal A.;
 RT "Five novel mutations in the L1CAM gene in families with X linked
 hydrocephalus.";
 RL J. Med. Genet. 33:103-106(1996).
 RN [21]
 RP VARIANTS HSAS GLN-184; 439-VAL--THR-443 DEL; CYS-784 AND
 RX 936-LEU-LEU-948 DEL.
 RX MEDLINE=97338664; PubMed=9195224;
 RA Macfarlane J.R., Du J.-S., Peppe M.E., Ramsden S., Donnai D.,
 Charlton R., Garrett C., Tolmie J., Yates J.R.W., Berry C., Goudie D.,
 Moncla A., Lunt P., Hodgson S., Jouet M., Kenrick S.;
 RT "Nine novel L1 CAM mutations in families with X-linked
 hydrocephalus.";
 RL Hum. Mutat. 9:512-518(1997).
 RN [22]
 RP VARIANTS HSAS/MASA ASP-691; ARG-698 AND PRO-935.
 RX MEDLINE=98180721; PubMed=9521424;
 RA Du Y.-Z., Stivaeva A.K., Schwartz C.E.;
 RT "Multiple exon screening using restriction endonuclease
 fingerprinting (REF) detection of six novel mutations in the L1 cell
 adhesion molecule (L1CAM) gene.";
 RL Hum. Mol. Genet. 6:103-108(1997).

Query Match 8.34; Score 109; DB 1; Length 1257;
 Best Local Similarity 20.98; Pred. No. 1.6;
 Matches 46; Conservative 35; Mismatches 97; Indels 50; Gaps 9;

Oy	8	RHLTVL----	LOALPLPA--	ATGKNKVLKGGDVTCTASQKSIQFMKNSQIKI	61
Db	409	RHGLLLANAYIVVQ	PAKLITLADNQTVMVQGSTAYLLCKAGAPVPSQMLDEDTTV	468	
Oy	62	IGNQSFETKGP	KLNDRADSRSLMDQGNFPLIKNLKIEDSPYVCEVEDQKEEVQL	121	
Db	469	LDGERFP-----	-----	PNANGTIGIRLQANDGRYCCLANQNNVTIM	509
Oy	122	VFGITANSPTHLIQ-----	-----	QGSITLTLFS--PGSSPSVQCSPPRGKNIQ-----	164
Db	510	A-MLKVDATQITQGP	STIEKKGSRTVFCASFPSPLOPSTWRD--GRDIQELQDSD	567	
Oy	165	-----GGKTLVSQ	LELQSDSGTWTQVLQKQKVEFIDIVPAPASALPAP	209	
Db	568	KYFIEDGR-LVHSLD	SDGNTSCVASTELDVESRAQLLVGSGPQVP	616	

RESULT 38
 VGR2_MOUSE
 ID_VGR2_MOUSE STANDARD; PRT; 1367 AA.

AC p35918;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vascular endothelial growth factor receptor 2 precursor (EC 2.7.1.112)
 DE (VEGFR-2) (Protein-tyrosine kinase receptor flk-1) (Fetal liver kinase
 1) (kinase NKX)
 GN KDR OR FLK1 OR FLK-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Embryo;
 RX MEDLINE=93208880; Pubmed=7681362;
 RA Millaer B., Witzmann-Voos S., Schnurch H., Martinez R.,
 RA Meiller N.P.H., Risaau W., Ullrich A.;
 RT "High affinity VEGF binding and developmental expression suggest
 RT Flk-1 as a major regulator of vasculogenesis and angiogenesis.";
 RL Cell 72:835-846(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/He; TISSUE=Fetal liver;
 RX MEDLINE=92020984; Pubmed=1717995;
 RA Mathews W., Jordan C.T., Gavin M., Jenkins N.A., Copeland N.G.,
 RA Lemlichka I.R.;
 RT "A receptor tyrosine kinase cDNA isolated from a population of
 RT enriched primitive hematopoietic cells and exhibiting close genetic
 RT linkage to c-kit.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:9026-9030(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93141255; Pubmed=8423988;
 RA Oelrichs R.B., Reid H.H., Bernard O., Ziemecki A., Wilks A.F.;
 RT "VW/FLK-1: a putative receptor protein tyrosine kinase isolated from
 RT E10 embryonic neuroepithelium is expressed in endothelial cells of
 RT the developing embryo.";
 RL Oncogene 8:11-18(1993).
 RN [4]
 RP SEQUENCE OF 1-15 FROM N.A.
 RX MEDLINE=96032749; Pubmed=7559454;
 RA Patterson C., Petrella M.A., Hsieh C.-M., Yoshizumi M., Lee M.-E.,
 RA Harber E.;
 RT "Cloning and functional analysis of the promoter for KDR/flk-1, a
 RT receptor for vascular endothelial growth factor.";
 RL J. Biol. Chem. 270:23111-23118(1995).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=93361481; Pubmed=8356051;
 RA Quinn T.P., Peters K.G., de Vries C., Ferrara N., Williams L.T.;
 RT "Fetal liver kinase 1 is a receptor for vascular endothelial growth
 RT factor and is selectively expressed in vascular endothelium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7533-7537(1993).
 CC -1- FUNCTION: RECEPTOR FOR VEGF OR VEGF-C. HAS A TYROSINE-PROTEIN
 CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
 CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF
 CC VASCULAR PERMEABILITY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT HEART, LUNG,
 CC KIDNEY, BRAIN AND SKELETAL MUSCLE, BUT IS ALSO EXPRESSED AT LOWER
 CC LEVELS IN MOST OTHER ADULT TISSUES.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases.
 CC CSF-1/PDGF receptor subfamily.
 CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.
 CC HSSP; P11362; 1FGK.
 CC MGD; MGI:96683; Kdr.
 CC GO; GO:0045165; Pcell fate commitment; IMP.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003598; Ig_c2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001824; ReceptTykineIII.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; Ig_6.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 2.
 CC SMART; SM00408; Igc2; 1.
 CC SMART; SM00219; TyKc; 1.
 CC PROSITE; PS0835; IG_LIKE; 5.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 CC DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 CC KW Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;
 CC KW Transmembrane; Glycoprotein; Phosphoprotein; ATP-binding;
 CC Immunoglobulin domain; Repeat.
 CC FT CHAIN 1 19
 CC FT 1 1367
 CC FT 20 762
 CC FT 763 784
 CC FT 785 1367
 CC FT 146 111
 CC FT 143 209
 CC FT 226 325
 CC FT 330 416
 CC FT 423 542
 CC FT 549 656
 CC FT 665 751
 CC FT 832 1160
 CC FT 838 846
 CC FT 866 866
 CC FT 1026 1026
 CC FT 46 46
 CC FT 98 98
 CC FT 145 145
 CC FT 160 160
 CC FT 247 247
 CC FT 320 320
 CC FT 376 376
 CC FT 397 397
 CC FT 509 509
 CC FT 521 521
 CC FT 578 578
 CC FT 611 611
 CC FT 617 617
 CC FT 629 629
 CC FT 673 673
 CC FT 702 702
 CC FT 719 719
 CC FT 1057 1057
 CC FT 25 25
 CC FT 679 679
 CC FT 783 783
 CC FT 917 917
 CC FT 1341 1367
 CC FT 1367 AA; 152516 MM; EFC99704F1DCA266 CRC64;
 CC -----
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X70842; CAA50192.1; -
 CC DR EMBL; X59397; CAA42040.1; -
 CC DR EMBL; S53103; AAB35043.1; -
 CC DR EMBL; X89777; CAA61917.1; -
 CC DR PIR; A41228; A41228.

```

Query Match          8 3%; Score 109; DB 1; Length 1367;
Best Local Similarity 21.0%; Pred. No. 1.8;
Matches 45; Conservative 36; Mismatches 73; Indels 60; Gaps 9

Oy      9 HLLVLQALLPAATGKNKVVLGKKGDPTVELCTASQKSLIQPHKNSNQIKLNGQSF 68
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      544 HVINGELTVPQAAPTEQ-----ESVSLCTIDRTFENLYTK-----LSQSATS 590
Oy      69 LTKGPSKLNRADRSRSLW-----DQGNPLII--KNLKIEDSDTYICEVEDQKEE 117
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      591 VHMGES-LTPVCNKNDALMKLTNGTFNSSTNIDLIVARONSLQGGDYVCAQDKTKXK 649
Oy      118 -----VOLLVGLTNSDTHLQGSGSLTLESPPGSSPSYQCSPRKCN----- 162
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      650 RHCLVKQLILERMAMPITGMLENQTTI-----GETIEVTC--PAGNPPTHITWRD 701
Oy      163 -----IQGKTLSVSQLELOPSGTWC 184
         |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      702 NETLVEDSGIVLRDGNRLTIIRRVAKDEGLYTIC 735

RESULT 39
PGBM_HUMAN        STANDARD:      PRT:   4391 AA.
AC      P98160.Q16287.Q9H3V5;
DT      01-OCT-1996 (Rel. 34, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Basement membrane-specific heparan sulfate proteoglycan core
DE      protein precursor (HSPG) (perlecan) (PLC).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homindaes; Homo.
OX      NCBI_Taxid=9606;
[1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=92112994; PubMed=1730768;
RX      Kallunki P., Tryggvason K.;
RT      "Human basement membrane heparan sulfate proteoglycan core protein: a
RT      467-kD protein containing multiple domains resembling elements of the
RT      low density lipoprotein receptor, laminin, neural cell adhesion
RT      molecules, and epidermal growth factor."
RL      J. Cell Biol. 116:559-571(1992).
[2]
RN      SEQUENCE FROM N.A.
RP      TISSUE=Colon, and Skin;
RC      MEDLINE=92235084; PubMed=1569102;
RX      Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT      "Primary structure of the human heparan sulfate proteoglycan from
RT      basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT      domain homologous to the low density lipoprotein receptor, laminin,
RT      neural cell adhesion molecules, and epidermal growth factor."
RL      J. Biol. Chem. 267:8544-8557(1992).
[3]
RN      SEQUENCE OF 22-4391 FROM N.A., AND VARIANT SJ51 TYR-1532.
RP      MEDLINE=20553141; PubMed=1101850;
RX      Nicole S., Davolio C.-S., Topaloglu H., Cartolico L., Barral D.,
RA      Belgheon P., Ben-Hamida C., Hamouda H., Craud C., White P.S.,
RA      Samson D., Urtizberea J.A., Lehmann-Horn F., Weissenbach J.,
RA      Henlart F., Fontaine B.;
RT      "Perlecan, the major proteoglycan of basement membranes, is altered in
RT      patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia).";
RL      Nat. Genet. 26:480-483(2000).
[4]
RP      SEQUENCE OF 1016-1470 FROM N.A.
RC      TISSUE=Colon;
RX      MEDLINE=91365376; PubMed=1679749;
RA      Dodge G.F., Kovalevsky I., Chu M.L., Hassell J.R., McBride O.W.,
RA      Yi H.G., Iozzo R.V.;
RT      "Heparan sulfate proteoglycan of human colon: partial molecular
RT      cloning, cellular expression, and mapping of the gene (HSPG2) to the
RT      short arm of human chromosome 1."

```

RL Genomics 10:673-680(1991).
 RN [5]
 RP SEQUENCE OF 890-1396 FROM N.A.
 RX TISSUE=Fibroblasts;
 RC MEDLINE=92120660; PubMed=1685141;
 RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
 RA Trygvaeson K.;
 RA "Cloning of human heparan sulfate proteoglycan core protein,
 RT assignment of the gene (HSPG2) to 15p6.1-7p35 and identification of
 RT a BamHI restriction fragment length polymorphism.";
 RN Genomics 11:389-396(1991).
 RP SEQUENCE OF 1-21 FROM N.A.
 RX MEDLINE=94052171; PubMed=8234307;
 RA Cohen I.R., Graessell S., Murdoch A.D., Iozzo R.V.;
 RA "Structural characterization of the complete human perlecan gene and
 RT its promoter.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITE ASN-2121.
 RX MEDLINE=22660472; PubMed=12754519;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RA "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, scorable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666(2003).
 CC -1- FUNCTION: This protein is an integral component of basement
 CC membranes. It is responsible for the fixed negative electrostatic
 CC charge and is involved in the charge-selective ultrafiltration
 CC properties. It serves as an attachment substrate for cells.
 CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
 CC dimers or stellate structures. It interacts with other basement
 CC membrane components such as laminin, prolargin and collagen type
 CC IV.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Found in the basement membranes.
 CC -1- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
 CC AND O-LINKED OLIGOSACCHARIDES.
 CC -1- DISEASE: Defects in HSPG2 are the cause of Schwartz-Jampel
 CC syndrome (SSJ1) [MIM:255800]; a rare autosomal recessive disorder
 CC characterized by permanent myotonia (prolonged failure of muscle
 CC relaxation) and skeletal dysplasia, resulting in reduced stature,
 CC kyphoscoliosis, bowing of the diaphyses and irregular epiphyses.
 CC -1- SIMILARITY: Contains 4 LDL-receptor class A domains.
 CC -1- SIMILARITY: Contains 11 laminin EGF-like domains.
 CC -1- SIMILARITY: Contains 3 laminin IV domains.
 CC -1- SIMILARITY: Contains 22 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 4 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 SEA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, X62515; CAA44373.1; -;
 DR EMBL, M85289; AAA52700.1; -;
 DR EMBL, AL445795; CAC18534.1; -;
 DR EMBL, M64283; AAA52699.1; -;
 DR EMBL, S76436; AAB21121.2; -;
 DR EMBL, I22078; -; NOT_ANNOTATED_CDS.
 DR PIR, A38096; A38096.
 DR HSSP, P00740; 1EDM.
 DR Slens-2DPAGE; P98160; -;
 DR Genew; HGNC:5273; HSPG2.
 DR MIM; 142461; -;
 DR MIM; 255800; -;
 DR InterPro; IPR008985; ConA_like_1ec_g1.
 DR InterPro; IPR007742; EGF_2.
 DR InterPro; IPR006209; EGF-like.

DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003599; IG.
 DR InterPro: IPR003598; IG_c2.
 DR InterPro: IPR003596; IG_v.
 DR InterPro: IPR000034; Laminin_B.
 DR InterPro: IPR002048; Laminin_EGF.
 DR InterPro: IPR001791; Laminin_G.
 DR InterPro: IPR002172; LDL_receptor_A.
 DR InterPro: IPR000082; SEA_domain.
 DR Pfam: PF00008; EGF; 4.
 DR Pfam: PF00047; IG; 22.
 DR Pfam: PF00052; Laminin_B; 3.
 DR Pfam: PF00053; Laminin_EGF; 7.
 DR Pfam: PF00054; Laminin_G; 3.
 DR Pfam: PF00057; Idl_recept_a; 4.
 DR Pfam: PF01390; SEA; 1.
 DR PRINTS: PR00261; LDLRECEPTOR.
 DR ProDom: PD003031; Laminin_B; 3.
 DR SMART: SM00181; EGF; 15.
 DR SMART: SM00180; EGF_Lam; 12.
 DR SMART: SM00409; IG; 22.
 DR SMART: SM00408; IGC2; 21.
 DR SMART: SM00406; IG; 7.
 DR SMART: SM00281; Lamb; 3.
 DR SMART: SM00282; Lamb; 3.
 DR SMART: SM00192; LDLa; 4.
 DR PROSITE: PS00022; EGF_1; 9.
 DR PROSITE: PS01186; EGF_2; 6.
 DR PROSITE: PS50025; EGF_3; 4.
 DR PROSITE: PS50835; IG_LIKE; 22.
 DR PROSITE: PS50025; LAM_G_DOMAIN; 3.
 DR PROSITE: PS01248; LAMININ_TYPE_EGF; 11.
 DR PROSITE: PS01209; LDLRA_1; 4.
 DR PROSITE: PS50068; LDLRA_2; 4.
 DR PROSITE: PS50024; SEA; 1.
 KW Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
 KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
 KW Extracellular matrix; EGF-like domain; Disease mutation.
 FT SIGNAL 1 21
 FT CHAIN 22 4391
 FT DOMAIN 80 194
 FT DOMAIN 198 235
 FT DOMAIN 284 320
 FT DOMAIN 324 360
 FT DOMAIN 367 404
 FT DOMAIN 405 504
 FT DOMAIN 521 530
 FT DOMAIN 531 730
 FT DOMAIN 731 763
 FT DOMAIN 764 813
 FT DOMAIN 814 871
 FT DOMAIN 879 923
 FT DOMAIN 924 933
 FT DOMAIN 934 1125
 FT DOMAIN 1126 1158
 FT DOMAIN 1159 1208
 FT DOMAIN 1209 1265
 FT DOMAIN 1275 1324
 FT DOMAIN 1325 1334
 FT DOMAIN 1335 1529
 FT DOMAIN 1530 1562
 FT DOMAIN 1563 1612
 FT DOMAIN 1613 1670
 FT DOMAIN 1671 1771
 FT DOMAIN 1772 1865
 FT DOMAIN 1866 1955
 FT DOMAIN 1956 2051
 FT DOMAIN 2052 2151
 FT DOMAIN 2152 2244
 FT DOMAIN 2245 2340

FT DOMAIN 2341 2436 IG-LIKE C2-TYPE 9.
 FT DOMAIN 2437 2533 IG-LIKE C2-TYPE 10.
 FT DOMAIN 2534 2629 IG-LIKE C2-TYPE 11.
 FT DOMAIN 2630 2726 IG-LIKE C2-TYPE 12.
 Query Match 8.3%; Score 109; DB 1; Length 4391;
 Best local similarity 19.5%; Pred. No. 7.4; Indels 116; Gaps 11;
 Matches 63; Conservative 36; Mismatches 108; Indels 116; Gaps 11;
 QY 12 LVYQLALPAAQGNKVLGGKGDVETLTCTASQKKS1QFHWKN-----SNQI- 59
 DB 3291 ILIHVSPRYATVTPHSAVSQGETVQLQCLAHGPPPLTFQMSRVSSLPGRATANELL 3350
 QY 60 -----KILNQSS-----FLTKGP----- 73
 DB 3351 HPERAAPEDSGRYRCRTKTKVSAEFAQLLVQGGSLPATSPAGSPPTVQVTPOLET 3410
 QY 74 -----SKLNDRAISRSLMDQGNF-----LIINKLIDSDTYICEVE- 112
 DB 3411 KSLIGASVEFHCAVPSDRGTQLRMFKEGGQLPPGHSHQSDVLKAIQNLDQSCGGTYICQAHG 3470
 QY 113 -----DQKEVQLVFGLTANSPTLLQGSITLTLESPPGSSP 151
 DB 3471 PMGKAQASALVYQALPSVLINIRTSVQTVVG-----HAVERFCLALGPKPGVTVWS 3523
 QY 152 SVQCRSPRGKNIQGGKTLVSQLELDGSGTWCTVYQ-----NKKVEFKIDIVRASALP 207
 DB 3524 KYGGHLRPQ-IVQSGGVVRIAVELADAGVCRCTAAGTTQSHVLLVQALPOLIS-MP 3581
 QY 208 AP---PTGSALPDPQTASALPDP 227
 DB 3582 QEVVPAGSAVAFPCIASGYPFP 3604
 RESULT 40
 FCEA_HUMAN STANDARD; PRT; 257 AA.
 ID FCEA_HUMAN
 AC P12319;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE High affinity immunoglobulin epsilon receptor alpha-subunit precursor
 DE (FCER1) (IG Epsilon receptor, alpha-subunit) (Fc-epsilon RI-alpha).
 GN FCER1A OR FCE1A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88233953; PubMed=2967464;
 RA Kochan J., Pettine L.F., Hakimi J., Kishi K., Kinet J.-P.,
 RT "Isolation of the gene coding for the alpha subunit of the human high
 affinity IgE receptor."
 RL Nucleic Acids Res. 16:3584-3584(1988).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX TISSUE=Mast cells;
 RC MEDLINE=88158102; PubMed=2964640;
 RA Shimizu A., Tepler I., Benfey P.N., Berenstein E.H., Siraganian R.P.,
 RA Leder P.;
 RT "Human and rat mast cell high-affinity immunoglobulin E receptors:
 RT characterization of putative alpha-chain gene products."
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988).
 RN [3]
 RP 3D-STRUCTURE MODELING OF 26-197.
 RX MEDLINE=93113350; PubMed=1472946;
 RA Padlan E.A., Helm B.A.;
 RT "A modeling study of the alpha-subunit of human high-affinity
 receptor for immunoglobulin-E."
 RL Receptor 2:129-144(1992).
 CC -I- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
 AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC

CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
 CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
 CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
 CC ALSO INDICES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
 CC DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protease.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL; X06948; CAA30025.1; -;
 DR EMBL; J03605; AAA36204.1; -;
 DR EMBL; A21606; CAA01564.1; -;
 DR PIR; S00682; S00682.
 DR PDB; 1ALS; 27-FEB-95.
 DR PDB; 1ALT; 27-FEB-95.
 DR PDB; 1F2Q; 08-JUN-00.
 DR PDB; 1J86; 29-AUG-01.
 DR PDB; 1J87; 29-AUG-01.
 DR PDB; 1J88; 29-AUG-01.
 DR PDB; 1J89; 05-SEP-01.
 DR Genew; HGNC:3609; FCER1A.
 DR MIM; 147140; -;
 DR GO; GO:0005887; C:Integral to plasma membrane; TAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR Pfam; PF00047; Ig_2.
 DR SMART; SM00408; IGc2; 1.
 DR PROSITE; PS50835; IG_LIKE; 2.
 DR IGF-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
 KM Immunoglobulin domain; Repeat; 3D-structure.
 FT SIGNAL 1
 FT CHAIN 25
 FT 26 257 HIGH AFFINITY IMMUNOGLOBULIN EPSILON
 FT DOMAIN 26 205 RECEPTOR ALPHA-SUBUNIT.
 FT TRANSMEM 226 224 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 225 257 POTENTIAL.
 FT DOMAIN 30 110 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 111 193 IG-LIKE 1.
 FT DISULFID 51 93 IG-LIKE 2.
 FT CARBOHYD 46 67 BY SIMILARITY.
 FT DISULFID 132 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 46 67 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 67 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT STRAND 31 35
 FT HELIX 39 42
 FT TURN 43 43
 FT STRAND 48 54
 FT STRAND 61 66
 FT TURN 67 68
 FT STRAND 69 71
 FT STRAND 74 74
 FT TURN 75 76
 FT STRAND 77 79
 FT HELIX 82 87
 FT STRAND 91 96
 FT TURN 98 100
 FT HELIX 101 101
 FT STRAND 104 107
 FT STRAND 114 116
 FT HELIX 120 124
 FT STRAND 128 133

FT STRAND 143 145
 FT HELIX 146 148
 FT TURN 149 149
 FT STRAND 150 150
 FT TURN 152 153
 FT STRAND 154 155
 FT TURN 156 159
 FT STRAND 160 163
 FT HELIX 165 170
 FT STRAND 175 178
 FT STRAND 181 182
 FT TURN 183 185
 FT STRAND 186 187
 FT STRAND 190 193
 SQ SEQUENCE 257 AA; 29596 MW; F183BB2357DDAD58 CRC64;

Query Match 8.28; Score 108; DB 1; Length 257;
 Best Local Similarity 27.18; Pred. No. 0.27;
 Matches 55; Conservative 22; Mismatches 64; Indels 62; Gaps 10;

QY 10 LLLVLLQLALP---AATQGNKVYLQK-----KGDVVELTCTASQKKSIQFHKNKSNQI 59
 DB 10 LLLVLLFPADGVLAVPQKVSINPNNRIFKGENVTITC----- 51
 QY 60 KILGNQSFLLTKGPKINDRADSRSLMDQGNFPLIKNLKIEDSDTYICEVE--DQKEE 117
 DB 52 ---NGNNFFPVSTKPFHNG---SLSEETNSLNIYNNAKFEBSGKCYKQHQVNESER 103
 QY 118 VOLLVPLGILTNSDTHLLQGSLLTLLSPGSSPSVQCRSPRGKNI-----QGGKTL-- 169
 DB 104 VYLEVF-----SDLLLLQA-SAEVMEGQP---LFLRCHGRMNDVYKVIYKDGEMALKY 154
 QY 170 -----SVSOLHELDSDGTWTCT 185
 DB 155 WYENHNISTNATVEDSGTYTCT 177

RESULT 41
 A33_HUMAN STANDARD; PRT; 319 AA.
 ID A33_HUMAN
 AC Q99795;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Cell surface A33 antigen precursor (Glycoprotein A33).
 GN GPA33.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Colon carcinoma;
 RX MEDLINE=97165045; PubMed=9012807;
 RA Heath J.K., White S.J., Johnstone C.N., Catimel B., Simpson R.J.,
 RA Moritz A.L., Tu G.-F., Johnstone C.N., Whitehead R.H., Green L.C.,
 RA Scott A.W., Ritter G., Cohen L., White S., Old L.J., Nice E.C.,
 RA Burgess A.W.;
 RA "The human A33 antigen is a transmembrane glycoprotein and a novel
 RT member of the immunoglobulin superfamily.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:469-474(1997).
 RP POST-TRANSLATIONAL MODIFICATIONS.
 RX MEDLINE=9736159; PubMed=9245713;
 RA Ritter G., Cohen L.S., Nice E.C., Catimel B., Burgess A.W.,
 RA Moritz R.L., Ji H., Heath J.K., White S.J., Welt S., Old L.J.,
 RA Simpson R.J.;
 RT "Characterization of posttranslational modifications of human A33
 RT antigen, a novel palmitoylated surface glycoprotein of human
 RT gastrointestinal epithelium.";
 CC Biochem. Biophys. Res. Commun. 236:682-686(1997).
 CC -1- FUNCTION: May play a role in cell-cell recognition and signaling.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.


```

-1- TISSUE SPECIFICITY: Expressed in normal gastrointestinal
CC epithelium and in 95% of colon cancers.
CC
CC -1- PPM: N-GLYCOSYLATED, CONTAINS APPROXIMATELY 8 KDA OF N-LINKED
CC CARBOHYDRATE.
CC
CC -1- PPM: Palmitoylated.
CC
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: U79725; AAC50957.1; -.
CC
CC Genew: HGNC:4445; GPA33.
CC
CC MIM: 602171; -.
CC
CC GO: GO:0005888; C:proteoglycan integral to plasma membrane; TAS.
CC
CC GO: GO:0004872; F:receptor activity; TAS.
CC
CC InterPro: IPR007110; IG-like.
CC
CC InterPro: IPR003596; IG_V.
CC
CC Pfam: PF00047; Ig; 2.
CC
CC SMART: SMO0406; IGV: 1.
CC
CC PROSITE, PS50835; IG-LIKE; 2.
CC
CC Immunoglobulin domain; Lipoprotein; Palmitate; Glycoprotein;
CC Transmembrane; Signal; Antigen.
CC
CC FT SIGNAL 1 21
CC FT CHAIN 22 319 CELL SURFACE A33 ANTIGEN.
CC FT DOMAIN 22 235 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 236 256 POTENTIAL.
CC FT DOMAIN 257 319 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 22 134 IG-LIKE V-TYPE.
CC FT DOMAIN 140 227 IG-LIKE C2-TYPE.
CC FT DOMAIN 258 261 POLY-CYS.
CC FT DISULFID 43 117 POTENTIAL.
CC FT DISULFID 146 222 POTENTIAL.
CC FT DISULFID 162 211 POTENTIAL.
CC FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC
CC SQ SEQUENCE 319 AA; 35632 MW; 9BRC7AFA5C2408E CRC64;

```

DT	15-JUL-1998	(Rel. 36, Created)
DT	15-JUL-1998	(Rel. 36, Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)
DE	Neural cell adhesion molecule 2 precursor (N-CAM 2).	
GN	NCAAM2 OR NCAM21.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
CX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Brain;	
RX	MEDLINE=97369930; PubMed=9226371.	
RA	Padoni-Giacobino A., Chen H., Antonarakis S.E.;	
RT	"Cloning of a novel human neural cell adhesion molecule gene (NCAM2)	
RT	that maps to chromosome region 21q21 and is potentially involved in	
RL	Down syndrome.";	
RN	Genomics 43:43-51(1997).	
RN	[2]	
RP	CARBOHYDRATE-LINKAGE SITES ASN-445 AND ASN-562.	
RX	MEDLINE=22660472; PubMed=12754519;	
RA	Zhang H., Li X.-J., Martin D.B., Aebersold R.;	
RT	"Identification and quantification of N-linked glycoproteins using	
RT	hydrazone chemistry, stable isotope labeling and mass spectrometry.";	
RL	Nat. Biotechnol. 21:660-666(2003).	
CC	-I- FUNCTION: May play important roles in selective fasciculation and	
CC	zone-to-zone projection of the primary olfactory axons.	
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.	
CC	-I- TISSUE SPECIFICITY: Expressed most strongly in adult and fetal	
CC	brain.	
CC	-I- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.	
CC	-I- SIMILARITY: Contains 2 fibronectin type III domains.	

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@lsb-sib.ch).	
CC	-----	
DR	EMBL: U75330; AAB80803.1; --	
DR	GeneW: HGNC:7657; NCAM2.	
DR	MIM: 602040; --	
DR	GO: GO:0016021; C:integral to membrane; TAS.	
DR	GO: GO:0005886; C:plasma membrane; TAS.	
DR	GO: GO:0007158; P:neuronal cell adhesion; TAS.	
DR	InterPro: IPRO06957; FN_III-like.	
DR	InterPro: IPRO03961; FN_III.	
DR	InterPro: IPRO07110; Ig_II-like.	
DR	InterPro: IPRO03598; Ig_C2.	
DR	Pfam: PF00041; fn3; 2.	
DR	Pfam: PF00047; Ig; 5.	
DR	SMART: SMO0060; FN3; 2.	
DR	SMART: SMO0408; IGc2; 5.	
DR	PROSITE: PS50835; IG_LIKE; 5.	
KM	Cell adhesion; Transmembrane;	
KM	Immunoglobulin domain; Signal.	
FT	SIGNAL	1..19
FT	CHAIN	20..637
FT	DOMAIN	20..697
FT	TRANSMEM	698..718
FT	DOMAIN	719..837
FT	DOMAIN	21..108
FT	DOMAIN	113..202
FT	DOMAIN	208..297
FT	DOMAIN	302..396
FT	DOMAIN	401..491
FT	DOMAIN	482..581
FT	DOMAIN	594..678
FT	DISULFID	42..93
FT	DISULFID	136..186
FT	DISULFID	232..281

```

FT DISULFID 322 380 PROBABLE.
FT CARBOHYD 422 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 562 562 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 837 AA; 92932 MW; C3D034106C5741C1 CRC64;

Query Match 8.2%; Score 108; DB 1; Length 837;
Best Local Similarity 19.0%; Pred. No. 1.1;
Matches 45; Conservative 43; Mismatches 85; Indels 64; Gaps 11;

QY 3 RG-VPRHLVLVLALLPAATQGNKVVYLGKGDVETLCTASOKKSIOFMKNSNOIKI 61
DB 193 RGEIDFRDIYIVNVPALISMPQKSFNTAERGEMTSCASGSPDALISMPNNG-KL 250
QY 62 LGNGSFLTKGPSKLNDAADSRSLMDGNFPLIKNKIKEDSDTYICEV-----EDQKE 116
DB 251 IEENEKYLKQ-----SNTETVTRNIIINSDDGPGYVCRAATNKAQBDEKQ 293
QY 117 E-VOLVFGLTANSDTHLQ-----GSLVLTLESPPGSSPSVQC----- 155
DB 294 AFLQVFN-----QPHILQLKNETTYNGQ-VTLVCAEGEPIPEITWKRAVNGFTTE 345
QY 156 --RSPRG---KNIQGKTLVSQLELDGSGTWTCTVLQ---NOKKVEFKIDIVPR 202
DB 346 GDKSPDRGIEVKGQGGSSSLHKIVKJLGSGRYDEAASRIQGHQKSYLDIEVAPK 402

RESULT 43
NC22 HUMAN STANDARD; PRT; 761 AA.
AC P13592; P13593;
ID NC22 HUMAN STANDARD; PRT; 761 AA.
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120)
DE (NCAM1-120) (CD56 antigen).
GN NCAM1 OR NCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM N-CAM 120).
RC TISSUE=Sketelal muscle;
RA MEDLINE=89305258; PubMed=3253057;
RA Barton C.H., Dickson G., Gower H.J., Rowett L.H., Putt W.,
RA Elsom V., Moore S.E., Goidis C., Walsh F.S.;
RT "Complete sequence and in vitro expression of a tissue-specific
RT phosphatidylinositol-linked N-CAM isoform from skeletal muscle.";
RL Development 104:165-173(1988).
RN [2]
RP SEQUENCE OF 491-761 FROM N.A. (ISOFORM N-CAM 120).
RC TISSUE=Sketelal muscle;
RA MEDLINE=87301755; PubMed=2887295;
RA Dickson G., Gower H.J., Barton C.H., Prentice H.M., Elsom V.L.,
RA Moore S.E., Cox R.D., Quinn C., Putt W., Walsh F.S.;
RT "Human muscle neural cell adhesion molecule (N-CAM): identification
RT of a muscle-specific sequence in the extracellular domain.";
RL Cell 50:1119-1130(1987).
RN [3]
RP SEQUENCE OF 491-655 FROM N.A. (ISOFORM C).
RC MEDLINE=89077552; PubMed=3203385;
RA Gower H.J., Barton C.H., Elsom V.L., Thompson J., Moore S.E.,
RA Dickson G., Walsh F.S.;
RT "Alternative splicing generates a secreted form of N-CAM in muscle
RT and brain.";
RL Cell 55:955-964(1988).

```

```

CC -1- FUNCTION: This protein is a cell adhesion molecule involved in
CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=N-CAM 120;
CC IsoId=P13592-2; Sequence=Displayed;
CC Name=N-CAM 140;
CC IsoId=P13591-1; Sequence=External;
CC Name=C; Synonyms=Secreted;
CC IsoId=P13592-1; Sequence=VSP 002587;
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC -1- DATABASE: NMBE=PRO; NOTE=CD guide CD56 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/row/cd/cd56.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16841; CA34739.1; -
DR EMBL; M17409; AA55912.1; -
DR EMBL; M22094; AA55910.1; -
DR EMBL; M22092; AA55911.1; -
DR EMBL; M22091; AA55911.1; JOINED.
DR PIR; A31635; A31635.
DR PIR; S07784; ITHUNG.
DR Genew; HGNC; 7656; NCAM1.
DR MIM; 116930; -
DR GO; GO:0016021; C: integral to membrane; TAS.
DR GO; GO:0005886; C: plasma membrane; TAS.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; fn3; 2.
DR Pfam; PF00047; fn3; 2.
DR SMART; SM00408; FN3; 2.
DR PROSITE; PS50835; IG_LIKE; 5.
DR Immunoglobulin domain; Cell adhesion; Glycoprotein; Repeat; Signal;
KW GPI-anchor; Alternative splicing.
FT FT SIGNAL 1 19
FT CHAIN 20 761
FT FT
FT DOMAIN 20 111
FT FT
FT DOMAIN 116 205
FT FT
FT DOMAIN 212 301
FT FT
FT DOMAIN 308 403
FT FT
FT DOMAIN 406 491
FT FT
FT DOMAIN 518 595
FT FT
FT DOMAIN 660 727
FT FT
FT DISULFID 41 96
FT DISULFID 139 189
FT DISULFID 235 287
FT DISULFID 329 385
FT DISULFID 426 479
FT CARBOHYD 222 222
FT CARBOHYD 315 315
FT CARBOHYD 347 347
FT CARBOHYD 423 423
FT CARBOHYD 449 449
FT CARBOHYD 478 478
FT VARSPLIC 635 655
SQ SEQUENCE 761 AA; 83770 MW; F0CAD3292AD7AB67E CRC64;

```

Query Match 8.1%; Score 107; DB 1; Length 761;
 Best Local Similarity 19.3%; Pred. No. 1.2; Indels 44; Gaps 7;
 Matches 44; Conservative 38; Mismatches 102; Indels 44; Gaps 7;
 3 RG-VPRFHLVLTQALLPAATGKRVVLGKGDVTELTCTASQKSIQPHWKNQDIKI 61
 196 RGEINFRDQIVNVPTTQARQNIYVATNANLQGSVTLVCDAGFEPPTMSW----- 247
 62 LGNGQSFLLTKGPKSLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVE---DQKEE 117
 248 -----TKDGEQIEQEEDDEKXIFSDSSQLTIKKVDKNDEAEYICIAENKAGQDAT 299
 118 VOLLVFG---LTANSTHLLQGSVLTLESPPGSSPSVQCR-----SPRKNIGCG- 166
 300 IHLKVFAPKPIITYVENQTAELEBQVTLTCEASGDPISPIITWTSTRNISSEKTLIDGHM 359
 167 -----KTLVSQLELQDSGWTG---TVLNCKKVEFKDIDYVR 202
 360 VVSHARVSSLTLSKSIQYTAGETICTASNTIQDSSGMYLEQVAPK 407
 RESULT 44
 NCAL HUMAN STANDARD; PRT; 848 AA.
 ID NCAL HUMAN
 AC P13591: Q15829; Q16180;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neural cell adhesion molecule 1, 140 kDa isoform precursor (N-CAM 140)
 DE (NCAM-140) (CD56 antigen).
 GN NCAM1 OR NCAM.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94356433; PubMed=8075973;
 RA Saito S., Taniio Y., Tachibana I., Hayashi S., Kishimoto T., Kawase I.;
 RT "Complementary DNA sequence encoding the major neural cell adhesion
 RT molecule isoform in a human small cell lung cancer cell line.";
 RL Lung Cancer 10:307-316(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91250739; PubMed=1710251;
 RA Lanier L.B., Chang C., Azuma M., Ruitenberg J.J., Hemperly J.J.,
 RA Phillips J.H.;
 RT "Molecular and functional analysis of human natural killer cell-
 RT associated neural cell adhesion molecule (N-CAM/CD56).";
 RL J. Immunol. 146:4421-4426(1991).
 RN [3]
 RP SEQUENCE OF 491-848 FROM N.A.
 RX MEDLINE=87301755; PubMed=2887295;
 RA Dickson G., Gower H.J., Barton C.H., Prentice H.M., Elsom V.L.,
 RA Moore S.E., Cox R.D., Quinn C., Purr W., Walsh F.S.;
 RT "Human muscle neural cell adhesion molecule (N-CAM): Identification
 RT of a muscle-specific sequence in the extracellular domain.";
 RL Cell 50:1119-1130(1987).
 CC -1- FUNCTION: This protein is a cell adhesion molecule involved in
 CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
 CC neurites, etc.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=N-CAM 140;
 CC IsoId=P13591-1; Sequence=Displayed;
 CC Name=N-CAM 120;
 CC IsoId=P13592-2; Sequence=External;
 CC Name=C; Synonym=Secreted;
 CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
 CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD56 entry;

CC MW="http://www.ncbi.nlm.nih.gov/prov/cd/cd56.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: S71824; AAB31836.1; -
 CC EMBL: U63041; AAB04558.1; -
 CC EMBL: M7410; AAB59913.1; -
 CC HSSP: P40189; 1BOU.
 CC Genew: HGNC:7656; NCAM1.
 CC MIM: 116930; -
 CC GO: GO:0016021; C:integral to membrane; TAS.
 CC GO: GO:0005886; C:plasma membrane; TAS.
 CC InterPro: IPR008957; FN III-like.
 CC InterPro: IPR003961; FN III.
 CC InterPro: IPR007110; Ig-like.
 CC InterPro: IPR003598; Ig_C2.
 CC Pfam: PF00041; fn3; 2.
 CC Pfam: PF00047; Ig; 5.
 CC SMART: SM00060; FN3; 2.
 CC SMART: SM00408; IGC2; 5.
 CC PROSITE: PS50835; Ig_LIKE; 5.
 CC KIMMUNOGLOBULIN domain; Cell adhesion; Glycoprotein; Repeat; Signal;
 CC Transmembrane; Alternative splicing.
 CC FT SIGNAL 1 19
 CC FT CHAIN 20 848
 CC FT DOMAIN 20 708
 CC FT TRANSMEM 709 729
 CC FT DOMAIN 730 848
 CC FT DOMAIN 20 111
 CC FT DOMAIN 116 205
 CC FT DOMAIN 212 301
 CC FT DOMAIN 308 403
 CC FT DOMAIN 406 491
 CC FT DOMAIN 518 595
 CC FT DOMAIN 660 727
 CC FT DISULFID 41 96
 CC FT DISULFID 139 189
 CC FT DISULFID 235 287
 CC FT DISULFID 329 385
 CC FT DISULFID 426 479
 CC FT CARBOHYD 222 222
 CC FT CARBOHYD 315 315
 CC FT CARBOHYD 347 347
 CC FT CARBOHYD 423 423
 CC FT CARBOHYD 449 449
 CC FT CARBOHYD 478 478
 CC FT CARBOHYD 215 215
 CC FT CONFLICT 239 239
 CC FT CONFLICT 490 490
 CC FT CONFLICT 599 600
 CC FT CONFLICT 720 721
 CC FT CONFLICT 811 811
 CC SQ SEQUENCE 848 AA; 93360 MW; 68D2F0C06C1CZAD CRC64;
 Query Match 8.1%; Score 107; DB 1; Length 848;
 Best Local Similarity 19.3%; Pred. No. 1.4; Indels 44; Gaps 7;
 Matches 44; Conservative 38; Mismatches 102; Indels 44; Gaps 7;
 3 RG-VPRFHLVLTQALLPAATGKRVVLGKGDVTELTCTASQKSIQPHWKNQDIKI 61
 196 RGEINFRDQIVNVPTTQARQNIYVATNANLQGSVTLVCDAGFEPPTMSW----- 247
 62 LGNGQSFLLTKGPKSLNDRADSRSLMDQGNFPIIKNLKIEDSDTYICEVE---DQKEE 117
 248 -----TKDGEQIEQEEDDEKXIFSDSSQLTIKKVDKNDEAEYICIAENKAGQDAT 299

```

Oy 118 VOLVFG----LTANSPTHLLQGSLTTLSPSSPSVOCR-----SPRGNIGG- 166
Dy 300 IHLKVFAPKITYVENQTAAMEBEVTLTCEASGPISITMRTSTRNISSEKTLDDHM 359
Dy 167 -----KTLVSQLELDGSGTWTC---TVLQNKVFEKIDIVP 202
Dy 360 VVRSHARVSLTSLKSIQYTDAGEYICTASNTIGDSQSMYLEVOYAPK 407

RESULT 45
DCC_MOUSE STANDARD; PRT; 1447 AA.
ID DCC_MOUSE
AC P70211;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tumor suppressor protein DCC precursor.
GN DCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN (1)
RC SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RP STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=96112625; Pubmed=8570174;
RA Cooper H.M., Armes P., Brito J., Gad J., Wilks A.F.;
RT "Cloning of the mouse homologue of the deleted in colorectal cancer
RT gene (mdc) and its expression in the developing mouse embryo.",
RL Oncogene 11:2243-2254(1995).
RN (2)
RP REVISIONS.
RC STRAIN=BALB/c; TISSUE=Brain;
RA Cooper H.M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=P70211-1; Sequence=Displayed;
CC Note=Isoform B is produced by alternative initiation at Met-85
CC of isoform A;
CC Name=C;
CC IsoId=P70211-2; Sequence=VSP_002501;
CC Event=Alternative initiation;
CC Comment=2 isoforms, A (shown here) and B, are produced by
CC alternative initiation at Met-1 and Met-85;
CC -1- TISSUE SPECIFICITY: In the embryo, expressed at high levels in the
CC developing brain and neural tube. In adult, highly expressed in
CC brain with very low levels found in testis, heart and thymus.
CC Isoform C is expressed only in the embryo.
CC -1- DEVELOPMENTAL STAGE: Low levels in early gestation. Highest levels
CC expressed during mid gestation. Levels decrease in late gestation
CC and remain at this level in the adult.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. DCC family.
CC -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 6 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X85788; CAA59786.1; -.
CC HSP: P56276; ITK.
CC MGI: MGI:94869; DCC.
CC InterPro: IPR008957; FN_III-1like.
CC InterPro: IPR003961; FN_III.
CC InterPro: IPR003962; FN_III_subd.

```

```

DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_c2.
DR Pfam: PF00041; fn3; 6.
DR Pfam: PF00047; Ig; 4.
DR PRINTS: PR00014; FNTYPEIII.
DR SMART: SM00060; FN3; 6.
DR SMART: SM00408; IGc2; 3.
DR PROSITE: PS50835; IG_LIKE; 4.
KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
KW Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1447
FT CHAIN 85 1447
FT INIT MET 85 85
FT DOMAIN 26 1097
FT TRANSMEM 1098 1122
FT DOMAIN 1123 1447
FT DOMAIN 36 135
FT DOMAIN 139 229
FT DOMAIN 234 326
FT DOMAIN 331 416
FT DOMAIN 426 522
FT DOMAIN 525 618
FT DOMAIN 619 716
FT DOMAIN 722 816
FT DOMAIN 840 940
FT DOMAIN 941 1042
FT DISULFID 61 117
FT DISULFID 161 212
FT DISULFID 261 310
FT DISULFID 352 400
FT CARBOHYD 60 60
FT CARBOHYD 94 94
FT CARBOHYD 299 299
FT CARBOHYD 318 318
FT CARBOHYD 478 478
FT CARBOHYD 628 628
FT CARBOHYD 702 702
FT VASPLIC 819 838
SQ SEQUENCE 1447 AA; 158298 MW; 0D1F1097C2D5B9F CRC64;

Query Match 8.1%; Score 107; DB 1; Length 1447;
Best Local Similarity 24.1%; Pred. No. 2.7; Mismatches 75; Indels 44; Gaps 8;
Matches 46; Conservative 26;

Oy 34 GDTVELTCTASQKKSIOFHWNNSQIKITLGNQGSFLTKGPSKLD-RADSRRLMDQGNF 92
Dy 154 GDTVLACVIGEMPTIHQKNQ-----DLNPLPGDSRVVVLPGS-- 195
Oy 93 PLIKNLKIEDSDTYICEVED-----QKEEVOLL-----VFGLTANSPTHLLQGG 137
Dy 196 ALQISRLQPGDSGYRCASNPASIRTNNAEVAILLDPGLHROLYFLQSPSNVIALEGK 255
Oy 138 SLTL---TLSPSS-----SPSVQCRSPRGNIGGKTLVSQLELDGSGTWTCVTLQ 188
Dy 256 DAILECCVGGPPSPFTWLAGEVYQLRSKK-YSLGSSNLLISNVTDSDSGITTCVYTY 314
Oy 189 NQKVEFEKIDI 199
Dy 315 KNEIISASDEL 325

RESULT 46
PGBM_MOUSE STANDARD; PRT; 3707 AA.
ID PGBM_MOUSE
AC 005793;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PGC).
GN HSPG2.

```

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RX MEDLINE=92078153; Pubmed=1744087;
 RA Noonan D.M., Fullie A., Valente P., Cai S., Horigan E., Sasaki M.,
 RT Yamada Y., Hassell J.R.;
 RT "The complete sequence of perlecan, a basement membrane heparan
 sulfate proteoglycan, reveals extensive similarity with laminin A
 chain, low density lipoprotein-receptor, and the neural cell adhesion
 molecule";
 RL J. Biol. Chem. 266:22939-22947 (1991).
 [2]
 RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=89034110; Pubmed=2972708;
 RA Noonan D.M., Horigan E.A., Ledbetter S.R., Vogel J.G., Sasaki M.,
 RT Yamada Y., Hassell J.R.;
 RT "Identification of cDNA clones encoding different domains of the
 basement membrane heparan sulfate proteoglycan";
 RL J. Biol. Chem. 263:16379-16387 (1988).
 CC -1- FUNCTION: This protein is an integral component of basement
 CC membranes. It is responsible for the fixed negative electrostatic
 CC charge and is involved in the charge-selective ultrafiltration
 CC properties. It serves as an attachment substrate for cells.
 CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
 CC dimers or stellate structures. It interacts with other basement
 CC membrane components such as laminin, prolargin and collagen type
 CC IV.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: Found in the basement membranes.
 CC -1- PFM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
 CC AND O-LINKED OLIGOSACCHARIDES.
 CC -1- SIMILARITY: Contains 4 LDL-receptor class A domains.
 CC -1- SIMILARITY: Contains 11 laminin EGF-like domains.
 CC -1- SIMILARITY: Contains 3 laminin IV domains.
 CC -1- SIMILARITY: Contains 15 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 3 laminin G-like domains.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -1- SIMILARITY: Contains 1 SEA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; M7174; AAA3911.1; -;
 DR EMBL; J04054; AAA3899.1; -;
 DR EMBL; J04055; AAA3912.1; -;
 DR PIR; S18252; S18252.
 DR PDB; 1GL4; 28-NOV-01.
 DR MGD; MGI:96257; Hsp92.
 DR GO; GO:0005604; C:basement membrane; IDA.
 DR GO; GO:0008104; P:protein localization; IMP.
 DR InterPro; IPR008985; ConA_like_1ec_g1.
 DR InterPro; IPR00742; EGF_2.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR000034; laminin_B.
 DR InterPro; IPR002049; laminin_EGF.
 DR InterPro; IPR001791; laminin_G.
 DR InterPro; IPR002172; LDL_receptor_A.
 DR InterPro; IPR000082; SEA_domain.
 DR Pfam; PF00008; EGF_4.
 DR Pfam; PF00047; IG_15.
 DR Pfam; PF00053; laminin_B; 3.
 DR Pfam; PF00053; laminin_EGF; 7.

DR Pfam; PF00054; laminin_G; 3.
 DR Pfam; PF00057; LDL_recept_a; 4.
 DR Pfam; PF01390; SEA; 1.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR PRODOM; PD003031; laminin_B; 3.
 DR SMART; SM00180; EGF_Lam; 7.
 DR SMART; SM00408; IGC2; 14.
 DR SMART; SM00281; Lamb; 3.
 DR SMART; SM00282; LamG; 3.
 DR SMART; SM00192; LDLa; 4.
 DR SMART; SM00200; SEA; 1.
 DR PROSITE; PS00022; EGF_1; 8.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01186; EGF_3; 4.
 DR PROSITE; PS00026; EGF_3; 4.
 DR PROSITE; PS00835; IG-LIKE; 15.
 DR PROSITE; PS00825; LAM_G_DOMAIN; 3.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
 DR PROSITE; PS01209; LDLRA_1; 4.
 DR PROSITE; PS00068; LDLRA_2; 4.
 DR PROSITE; PS00024; SEA_1.
 DR Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
 DR Heparan sulfate; laminin EGF-like domain; Immunoglobulin domain;
 DR Extracellular matrix; EGF-like domain; 3D-structure.
 DR SIGNAL 1 21
 FT CHAIN 22 3707
 FT 80 194
 FT 195 234
 FT 281 319
 FT 320 359
 FT 360 403
 FT 404 504
 FT 521 530
 FT 531 730
 FT 731 763
 FT 764 813
 FT 814 871
 FT 879 923
 FT 924 933
 FT 934 1125
 FT 1126 1158
 FT 1159 1208
 FT 1209 1265
 FT 1275 1325
 FT 1325 1334
 FT 1335 1529
 FT 1530 1562
 FT 1563 1612
 FT 1613 1670
 FT 1677 1771
 FT 1772 1865
 FT 1866 1954
 FT 1955 2049
 FT 2049 2148
 FT 2149 2244
 FT 2245 2343
 FT 2344 2436
 FT 2437 2532
 FT 2533 2619
 FT 2620 2720
 FT 2721 2809
 FT 2810 2895
 FT 2896 2980
 FT 2984 3162
 FT 3163 3241
 FT 3245 3425
 FT 3518 3705
 FT 3518 3705
 FT SITE 65 67
 FT SITE 71 73
 FT SITE 76 78
 FT SITE 3615 3617
 FT DISULFID 199 212
 BRASEMENT MEMBRANE-SPECIFIC HEPARAN
 SULFATE PROTEOGLYCAN CORE PROTEIN.
 SEA.
 LDL-RECEPTOR CLASS A 1.
 LDL-RECEPTOR CLASS A 2.
 LDL-RECEPTOR CLASS A 3.
 LDL-RECEPTOR CLASS A 4.
 IG-LIKE C2-TYPE 1.
 LAMININ EGF-LIKE 1 (N-TERMINAL).
 LAMININ EGF-LIKE 1 (C-TERMINAL).
 LAMININ EGF-LIKE 2.
 LAMININ EGF-LIKE 3.
 LAMININ EGF-LIKE 4 (INCOMPLETE).
 LAMININ EGF-LIKE 5 (N-TERMINAL).
 LAMININ EGF-LIKE 5 (C-TERMINAL).
 LAMININ EGF-LIKE 6.
 LAMININ EGF-LIKE 6 (C-TERMINAL).
 LAMININ EGF-LIKE 7.
 LAMININ EGF-LIKE 8.
 LAMININ EGF-LIKE 8 (N-TERMINAL).
 LAMININ EGF-LIKE 9 (N-TERMINAL).
 LAMININ EGF-LIKE 9 (C-TERMINAL).
 LAMININ EGF-LIKE 10.
 LAMININ EGF-LIKE 11.
 IG-LIKE C2-TYPE 2.
 IG-LIKE C2-TYPE 3.
 IG-LIKE C2-TYPE 4.
 IG-LIKE C2-TYPE 5.
 IG-LIKE C2-TYPE 6.
 IG-LIKE C2-TYPE 7.
 IG-LIKE C2-TYPE 8.
 IG-LIKE C2-TYPE 9.
 IG-LIKE C2-TYPE 10.
 IG-LIKE C2-TYPE 11.
 IG-LIKE C2-TYPE 12.
 IG-LIKE C2-TYPE 13.
 IG-LIKE C2-TYPE 14.
 IG-LIKE C2-TYPE 15.
 LAMININ G-LIKE 1.
 EGF-LIKE.
 LAMININ G-LIKE 2.
 LAMININ G-LIKE 3.
 HEPARAN SULFATE (POTENTIAL).
 HEPARAN SULFATE (POTENTIAL).
 HEPARAN SULFATE (POTENTIAL).
 MEDIATES MOTOR NEURON ATTACHMENT
 (POTENTIAL).
 BY SIMILARITY.

```

FT DISULFID 206 225 BY SIMILARITY.
FT DISULFID 219 234 BY SIMILARITY.
FT DISULFID 285 297 BY SIMILARITY.
FT DISULFID 292 310 BY SIMILARITY.
FT DISULFID 304 319 BY SIMILARITY.
FT DISULFID 325 337 BY SIMILARITY.
FT DISULFID 332 350 BY SIMILARITY.
FT DISULFID 344 359 BY SIMILARITY.
FT DISULFID 368 381 BY SIMILARITY.
FT DISULFID 375 394 BY SIMILARITY.
FT DISULFID 388 403 BY SIMILARITY.
FT DISULFID 428 479 BY SIMILARITY.
FT DISULFID 764 773 BY SIMILARITY.
FT DISULFID 766 780 BY SIMILARITY.
FT DISULFID 783 792 BY SIMILARITY.
FT DISULFID 795 811 BY SIMILARITY.
FT DISULFID 814 829 BY SIMILARITY.
FT DISULFID 816 839 BY SIMILARITY.
FT DISULFID 842 851 BY SIMILARITY.
FT DISULFID 854 869 BY SIMILARITY.
FT DISULFID 1159 1168 BY SIMILARITY.
FT DISULFID 1161 1175 BY SIMILARITY.
FT DISULFID 1178 1187 BY SIMILARITY.
FT DISULFID 1190 1206 BY SIMILARITY.
FT DISULFID 1209 1224 BY SIMILARITY.
FT DISULFID 1211 1234 BY SIMILARITY.
FT DISULFID 1237 1246 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1275 1287 BY SIMILARITY.
FT DISULFID 1293 1293 BY SIMILARITY.
FT DISULFID 1295 1304 BY SIMILARITY.
FT DISULFID 1307 1322 BY SIMILARITY.
FT DISULFID 1563 1572 BY SIMILARITY.
FT DISULFID 1565 1579 BY SIMILARITY.
FT DISULFID 1582 1591 BY SIMILARITY.
FT DISULFID 1594 1610 BY SIMILARITY.
FT DISULFID 1613 1628 BY SIMILARITY.
FT DISULFID 1615 1638 BY SIMILARITY.
FT DISULFID 1641 1650 BY SIMILARITY.
FT DISULFID 1653 1668 BY SIMILARITY.
FT DISULFID 1792 1839 BY SIMILARITY.
FT DISULFID 1886 1932 BY SIMILARITY.
FT DISULFID 1976 2021 BY SIMILARITY.
FT DISULFID 2073 2118 BY SIMILARITY.
FT DISULFID 2170 2215 BY SIMILARITY.
FT DISULFID 2268 2313 BY SIMILARITY.

Query Match 8.1%; Score 107; DB 1; Length 3707;
Best Local Similarity 21.7%; Pred. No. 8.5;
Matches 55; Conservative 39; Mismatches 95; Indels 64; Gaps 10;

```

```

OPCM_RAT
ID OPCM_RAT STANDARD; PRT; 345 AA.
AC P32735; P32735; Q01653; Q01654;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Opioid binding protein/cell adhesion molecule precursor (OBCAM)
DE (Opioid-binding cell adhesion molecule) (OPCML).
GN OPCML OR OBCAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92347701; Pubmed=1339369;
RA Lippman D.A., Lee N.M., Loh H.H.;
RT "Opioid-binding cell adhesion molecule (OBCAM)-related clones from a
RT rat brain cDNA library.";
RL Gene 117:249-254(1992).
RN [2]
RP SEQUENCE OF 195-214, AND GPI-ANCHOR.
RX MEDLINE=95198094; Pubmed=7891157;
RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,
RA Salzer J.L.;
RT "Cloning of neurotrophin defines a new subfamily of differentially
RT expressed neutral cell adhesion molecules.";
RL J. Neurosci. 15:2141-2156(1995).
CC -1- FUNCTION: Binds opioids in the presence of acidic lipids; probably
CC involved in cell contact.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=P32736-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P32736-2; Sequence=VSP_002612;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IgLON
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M88710; AAA40859.1; -
CC EMBL; M88711; AAA40860.1; -
CC EMBL; M88709; AAA40858.1; -
CC PIR; JCI238; JCI238.
CC PIR; JCI239; JCI239.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00408; Igc2; 2.
CC PROSITE; PSS0835; IG_LIKE; 3.
KW Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
KW Repeat; Signal; Alternative splicing; Lipoprotein.
FT SIGNAL 27
FT CHAIN 1 322
FT 28 322
FT PROPEP 323 345
FT DOMAIN 39 126
FT DOMAIN 136 219
FT DOMAIN 223 310
FT DISULFID 57 115
FT DISULFID 157 202

```

```

FT DISULFID 244 296 POTENTIAL.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 322 322 GPI-anchor amidated asparagine
(Potential).
FT VARSPLIC 1 27 MGVCGLFLPWLKLVVSLFLFLVPT -> MYHPAYWIF
SATVATLFLP (in Isoform 2).
FT FT /FtId=VSP_002612.
SQ SEQUENCE 345 AA; 38067 MW; A3181B0753F9658E CRC64;
Query Match 8.1%; Score 106.5; DB 1; Length 345;
Best Local Similarity 22.1%; Pred. No. 0.5;
Matches 58; Conservative 43; Mismatches 101; Indels 61; Gaps 10;
QY 11 LVVLQALFLPAATQGNKVVG-----KKGDVLELCTASQKKSIFHMKNSV 57
DB 14 LVVLSLKLFLVPTGVPRSGDATFPKAMDNVTVRQESATLCTDDRTV--KVAMLNRS 72
QY 58 QIKILGN-----QGSFLTGPSKLNDRASRSRLWDQNPFLIKNLKEDSDTYICE 110
DB 73 TLVAGNDKMSIDPRVILVNTPTQ-----YIMIQNVVDYEGPYTCS 116
QY 111 VE---DQKEVOLLVFG---LTNSDTHLLOGSLTLTLESPGSSPSVQCRSPRGKN 162
DB 117 VQVDNPKTSRVHLIVVPPQIMNISSDITVNEISSVTLCLAIQREPEVTVRHLISVKE 176
QY 163 IGG-----GKTLVSQLELDQSGMTCTVQLQ-----NOKKVEFIDIVPRASALPAPPTGS 213
DB 177 GGGFVSEDEYLETSDIRDSQGEYBCSALNDVAAPDKVKITVNPPTYS--KAKXNGV 234
QY 214 ALPDPQTASALPDPASALPAA 236
DB 235 SVGQKGLSC-----EASAVPMA 252
RESULT 48
DSCA HUMAN STANDARD; PRT; 2012 AA.
ID DSCA HUMAN 060469; 060468;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Down syndrome cell adhesion molecule precursor (CHD2).
GN DSCAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Brain;
RX MEDLINE=98087574; PubMed=9426258;
RA Yamakawa K., Huot Y.-K., Haendel M.A., Hubert R., Chen X.-N.,
RA Lyons G.B., Korenberg J.R.;
RT "DSCAM: a novel member of the immunoglobulin superfamily maps in a
RT Down syndrome region and is involved in the development of the
RT nervous system.";
RL Hum. Mol. Genet. 7:227-237(1998).
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=20384934; PubMed=10925149;
RA Aarawala K.L., Nakamura S., Teusum Y., Yamakawa K.;
RT "Down syndrome cell adhesion molecule DSCAM mediates homophilic
RT intercellular adhesion.";
RL Brain Res. Mol. Brain Res. 79:118-126(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,

```

```

RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA Soeda E., Onki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA Folley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W.,
RA Reinthal A., Kudoh J., Shibuya K., Kawasaki K., Aakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Sharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Kamber J., Beck A., Klages S., Hennig S., Rieseemann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- FUNCTION: CELL ADHESION MOLECULE THAT CAN MEDIATE CATION-
CC INDEPENDENT HOMOPHILIC BINDING ACTIVITY. COULD BE INVOLVED IN
CC NERVOUS SYSTEM DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PROBABLE). THE
CC SHORT ISOFORM MAY BE SECRETED.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long; Synonyms=CHD2-52;
CC IsoId=O60469-1; Sequence=Displayed;
CC Name=Short; Synonyms=CHD2-42;
CC IsoId=O60469-2; Sequence=VSP_002502, VSP_002503;
CC -1- TISSUE SPECIFICITY: Primarily expressed in brain.
CC -1- SIMILARITY: Contains 10 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 6 fibronectin type III domains.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF023450; AAC17967.1; -
CC EMBL: AF023449; AAC17966.1; -
CC EMBL: AF217525; AAF27525.1; -
CC EMBL: AL163283; CAB90464.1; -
CC EMBL: AL163282; CAB90436.1; -
CC EMBL: AL163281; CAB90444.1; -
CC Genew: HGNC:3039; DSCAM.
CC MIM: 602523; -
CC DR GO: 0005887; C: integral to plasma membrane; TAS.
CC DR GO: 0005624; C: membrane fraction; TAS.
CC DR GO: 0007155; P: cell adhesion; TAS.
CC DR GO: 0007399; P: neurogenesis; TAS.
CC DR InterPro: IPR008957; FN III-like.
CC DR InterPro: IPR003961; FN III.
CC DR InterPro: IPR007110; Ig-like.
CC DR InterPro: IPR003598; Ig_C2.
CC DR Pfam: PF00041; FN3; 6.
CC DR Pfam: PF00047; Ig; 9.
CC DR SMART: SM00408; FN3; 6.
CC DR SMART: SM00408; IGC2; 7.
CC DR PROSITE: PS50835; IG_LIKE; 9.
CC DR Immunoglobulin domain; Glycoprotein; Signal; Cell adhesion; Repeat;
CC Transmembrane; Alternative splicing.
CC KW Transmembrane; Signal; Cell adhesion; Repeat;
CC FT CHAIN 1 17
CC FT SIGNAL 18 2012
CC FT DOMAIN 18 1595
CC FT TRANSMEM 1596 1616
CC FT DOMAIN 1617 2012
CC FT DOMAIN 39 129
CC FT DOMAIN 125 216
CC FT DOMAIN 225 305
CC FT DOMAIN 313 401
CC FT DOMAIN 407 500
CC FT DOMAIN 504 592
CC FT DOMAIN 596 685
CC FT DOMAIN 690 783
CC FT DOMAIN 787 883
CC IG-LIKE C2-TYPE 9.

```



```

FT DOMAIN 885 972 FIBONECTIN TYPE-III 1.
FT DOMAIN 984 1076 FIBONECTIN TYPE-III 2.
FT DOMAIN 1088 1177 FIBONECTIN TYPE-III 3.
FT DOMAIN 1189 1273 FIBONECTIN TYPE-III 4.
FT DOMAIN 1285 1377 IG-LIKE C2-TYPE 10.
FT DOMAIN 1380 1463 FIBONECTIN TYPE-III 5.
FT DOMAIN 1477 1562 FIBONECTIN TYPE-III 6.
FT DISULFID 46 102 BY SIMILARITY.
FT DISULFID 145 197 BY SIMILARITY.
FT DISULFID 246 293 BY SIMILARITY.
FT DISULFID 335 385 BY SIMILARITY.
FT DISULFID 428 484 BY SIMILARITY.
FT DISULFID 525 575 BY SIMILARITY.
FT DISULFID 617 669 BY SIMILARITY.
FT DISULFID 711 766 BY SIMILARITY.
FT DISULFID 809 865 BY SIMILARITY.
FT DISULFID 1307 1359 BY SIMILARITY.
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 512 512 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 658 658 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 710 710 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 795 795 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 924 924 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1142 1142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1160 1160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1250 1250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1271 1271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1341 1341 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1488 1488 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VASPLIC 1562 1571 NFATL1NDGS -> KEARCKEFS (in isoform short).
FT VASPLIC 1572 2012 Missing (in isoform short).
FT CONFLICT 1893 2012 /FTid=VSP_002503.
FT HRPGLHLPLPYLPMFLNNGPQTSNLSGOACLEPOR
FT SRTGLRPTVLPPIEMEAASASSTREGSQMOPGAVATLPOR
FT EGAEIGQAKKSSQESLDSRGHLKGNKPNPAKSYTLV ->
FT IGVSTYCIHTEMTFC (IN REF. 1).
SQ SEQUENCE 2012 AA, 222259 MM; 0E3JCEB781A0834 CRC64;

Query Match 8.0%; Score 106; DB 1; Length 2012;
Best Local Similarity 21.5%; Pred. No. 4.8;
Matches 50; Conservative 34; Mismatches 75; Indels 74; Gaps 10;

OY 34 GDTVELTCTASQKSIQPHWKNQIKILGNQSFLLTKGPKKMDRADRSRLMDQGNFP 93
DB 239 GQVVELPCKALGHEPDYRMLKDNM-----PLELSGR-----FQKVTVG 277
OY 94 LIINKLEBDSPTICEVEDQKEVQLLVFGLTANSDTHLLOQSLTTLT-----ESPFG 148
DB 278 LILININPSDSGSIVCEVSNR-----YG-TAKVIGRLVVKPLAKTISPRVKSSVG 328
OY 149 SSPSVOCR-----SPRGKNIQ-----GKTLTSLVSOLELDQSGTWTCTV 186
DB 329 SSVLSSTSVGTEDQELSWYNGEILNP-GSNVRIITGINHMLMDHNVVSDGAYQCFV 387
OY 187 LONOKKVEFKIDIVPRASALPAPPTGSALPD--PQNASALDPDPPASALPAL 237
DB 388 RKDKLSADYVQVY-----LEDGTETKIIISAFSEKVSAPAEVSL 426

RESULT 49
NTRI_HUMAN STANDARD; PRT; 344 AA.
AC 09P121;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

```

```

DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurotrophin precursor (hnt).
GN NT...
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li G., Jin J., Tan X., Hu S., Yuan J., Qiang B.
RT "Cloning and identification of human neurotrophin full length cDNA.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Neural cell adhesion molecule.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IgLON
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF126426; AAF37591.1; -.
CC MIM; 607938; -.
CC DR GO; GO:0008038; P:neural cell recognition; TAS.
CC DR InterPro; IPR007110; Ig_Like.
CC DR InterPro; IPR003598; Ig_C.
CC DR Pfam; PF00047; Ig_3.
CC DR SMART; SM00408; IgC2; 2.
CC DR PROSITE; PS50835; IG_LIKE; 3.
CC DR Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor;
CC Repeat; Signal; Lipoprotein.
CC FT CHAIN 1 31 POTENTIAL.
CC FT PROPEP 32 321 NEUROTROPHIN.
CC FT DOMAIN 32 344 REMOVED IN MATURE FORM (POTENTIAL).
CC FT DOMAIN 136 218 IG-LIKE C2-TYPE 1.
CC FT DOMAIN 222 309 IG-LIKE C2-TYPE 2.
CC FT DISULFID 57 115 IG-LIKE C2-TYPE 3.
CC FT DISULFID 157 201 POTENTIAL.
CC FT DISULFID 243 295 POTENTIAL.
CC FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 70 70 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT LIPID 321 321 GPI-anchor amidated asparagine
CC (Potential).
SQ SEQUENCE 344 AA; 37971 MM; DA4D12C295ABBEA CRC64;

Query Match 8.0%; Score 105; DB 1; Length 344;
Best Local Similarity 20.8%; Pred. No. 0.65;
Matches 53; Conservative 48; Mismatches 108; Indels 46; Gaps 9;

OY 11 LVVLQALLPLPAALQGNKVVIG-----KGDVVELTCTASQKSIQPHWKN 57
DB 14 LVVLSLELLPLVPTGVVRSQDATFPKAMDNVTVRQESATLCTTDNRVT-RVAMLNRS 72
OY 58 QIKILGNQSFLLTKGPKKMDRADRSRLMDQGNFPILINLEKIBSDTYICEVEDQKEE 117
DB 73 TILYAGNDKWCCLDPVYVLNS-----TQYQYSEIQNVYVDEGPTCSVQDNHP 123
OY 118 -----VQLLVFGLTANSDTHLLOQSLTTLTSPGSSPSVOCR--SPRGKN-IOGG 166
DB 124 KTSRVHLIVQVSPKIVAISSDISINEGNISLCTIATGPREPTVTRHISPKAVGVSED 183
OY 167 KTLTSLVSOLELDQSGTWTCTVLO-----QKKVEFKIDIVPRASALPAPPTGSA 221

```

Db 184 EYEIIGITREOSGDEECASNDVAPVVRVTVNPPYIS----EAKGTGVPVGQKG 239
QY 222 SALPPPPASALPAA 236
Db 240 TLQCE---ASAVPSA 251

RESULT 50
NTRI_MOUSE STANDARD; PRT; 344 AA.
ID NTRI_MOUSE
AC 099PJO;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurotrophin precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Brain;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Kram T.H., Choi S.C., Kim J., Jeon U.W., Kim K.D., Lee S.H.;
RT "Cloning and expression of mouse neurotrophin gene in the developing
RT nervous system.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Eye;
RA MEDLINE=2388257; PubMed=12477932;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Krausberg R.L., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.C.,
RA Rana S.S., Loguailano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Neural cell adhesion molecule.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
CC family.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF282980; AAK00276.1; -
CC EMBL; BC023307; AAH23307.1; -
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC Pfam; PF00047; Ig_3.
CC SMART; SM00409; IG_3.
CC SMART; SM00408; IGc2; 3.

DR PROSITE; PSS0835; IG LIKE; 3.
KW Immunoglobulin domain, Cell adhesion, Glycoprotein, GPI-anchor;
KW Repeat; Signal; Lipoprotein.
FT SIGNAL 1 31
FT CHAIN 32 321
FT PROPEP 322 344
FT DOMAIN 39 126
FT DOMAIN 136 218
FT DOMAIN 222 309
FT DISULFID 57 115
FT DISULFID 157 201
FT DISULFID 243 295
FT CARBOHYD 44 44
FT CARBOHYD 70 70
FT CARBOHYD 152 152
FT CARBOHYD 284 284
FT CARBOHYD 292 292
FT CARBOHYD 305 305
FT CARBOHYD 321 321
FT LIPID 321 321
FT CONFLICT 75 75
FT CONFLICT 92 92
FT CONFLICT 119 119
FT CONFLICT 187 187
FT CONFLICT 213 213
FT CONFLICT 225 225
SQ SEQUENCE 344 AA; 37984 MW; C885BBA52C148554 CRC64;
Query Match 8.0%; Score 105; DB 1; Length 344;
Best local Similarity 20.8%; Pred. No. 0.65;
Matches 53; Conservative 48; Mismatches 108; Indels 46; Gaps 9;
QY 11 LVLQLALIPAAITQGNKVLG-----KKGDTVELTCTASQKKSIOFHKNKN 57
Db 14 LVVSLRLFLVPTGPVPSGDATEPKAMDNTVRSGBSATLRCTIDNRT- RVAMLNRS 72
QY 58 QKIIGNOSPLTKGSPKIDRADSRSLMDGNFLLIKNIKIEDSDYICEVEQKES 117
Db 73 TLVYANDKRCIDPRVLTLSN-----TQTSIEIQNDVVDDEGPTYSVQTDNHP 123
QY 118 -----VQLVFGLTNSDTHLLQGSITLTLESPPSSPSVOCR--SPRKN-IOGG 166
Db 124 KTSRVHLYQVSPKVEISDSISNGNNSITLCATGAPPEPTVTRHISPKAVGVSD 183
QY 167 KTLVSQGLQDSGTCTVLQN-----QKLYEPKIDYPRASALPAPRGSALPPQRTA 221
Db 184 EYEIIGITREOSGDEECASNDVAPVVRVTVNPPYIS----EAKGTGVPVGQKG 239
QY 222 SALPPPPASALPAA 236
Db 240 TLQCE---ASAVPSA 251

RESULT 51
NTRI_MOUSE STANDARD; PRT; 344 AA.
ID NTRI_MOUSE
AC Q62718;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neurotrophin precursor (GP65).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 217-229.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=95198094; PubMed=7891157;
RA Struyk A.F., Canoll P.D., Wolfgang M.J., Rosen C.L., D'Eustachio P.,
RA Salzer J.L.;
RT "Cloning of neurotrophin defines a new subfamily of differentially

Rt		expressed neural cell adhesion molecules.";
Ft	J.	Neurosci. 15;216(1995).
Cc	-1-	FUNCTION: Neural cell adhesion molecule.
Cc	-1-	SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
Cc	-1-	TISSUE SPECIFICITY: Central nervous system.
Cc	-1-	DEVELOPMENTAL STAGE: Expressed at high levels in several developing projection systems; in neurons of the thalamus, substantia nigra, and lower cortical laminae in the forebrain and in the pontine nucleus, cerebellar granule cells, and Purkinje cells in the hindbrain.
Cc	-1-	SIMILARITY: Belongs to the immunoglobulin superfamily. IgLON family.
Cc	-1-	SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
Cc	-1-	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Molecular Biology Laboratory. There are no restrictions on its reuse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).
Df	EMBL:	I56845; AAA67445.1; .
Df	PIR:	I56551; I56551.
Df	InterPro:	IPR007110; Ig-Like.
Df	InterPro:	IPR003598; Ig_C2.
Df	Pfam:	PF00047; Ig_3.
Dr	SMART:	SMO0408; IGc2; 2.
Dr	PROSITE:	PS00835; IG_LIKE; 3.
Kw	Immunoglobulin domain; Cell adhesion; Glycoprotein; GPI-anchor; Repeat; Signal; Lipoprotein.	
Ft	SIGNAL	1 31 POTENTIAL.
Ft	CHAIN	32 321 NEURORITMIN.
Ft	PROPEP	322 344 REMOVED IN MATURE FORM (POTENTIAL).
Ft	DOMAIN	39 126 IG-LIKE C2-TYPE 1.
Ft	DOMAIN	136 218 IG-LIKE C2-TYPE 2.
Ft	DOMAIN	222 309 IG-LIKE C2-TYPE 3.
Ft	DISULFID	57 115 POTENTIAL.
Ft	DISULFID	157 201 POTENTIAL.
Ft	DISULFID	243 295 POTENTIAL.
Ft	CARBOHYD	44 44 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CARBOHYD	70 70 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CARBOHYD	152 152 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CARBOHYD	216 216 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CARBOHYD	284 284 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CARBOHYD	292 292 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CARBOHYD	305 305 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	CARBOXYD	321 321 N-LINKED (GLCNAC...) (POTENTIAL).
Ft	Lipid	321 321 GPI-anchor amidated asparagine (Potential1).
SQ	SEQUENCE	344 AA; 37998 MW; CBH39BE5JB3B224 CRC64;
Query Match		Score 7.9%; DB 1; Length 344;
Best Local Similarity		20.8%; Pred.No. 0.77;
Matches	53;	Conservative 47; Mismatches 109; Indels 46; Gaps 9;
Oy	11	LVVLQLATLPATOGKRVLG-----KKGDVELLTASOKSIOFHMKSN 57 ::: :: Db LVVSLRLFLPVPGVPVRSGDATPRKAMDVTYVGGSALTICRTINDRYI-RVAAMLNRS 72 : Oy 58 QIKLGNCGSPFTPKPSKLNDRAUSRRSLMDQGFPILLIKULKIEDSDTYICEVDKEE 117 : Db 73 TLIVAGNWKWCLDPVVLLSN-----IQTVGISIEIGNVDVDEGPTCSVGTDNHP 123 - - - - VOLIVEPLTNSTDTHLLGGOSLTITLESPGSSPSVQC-R-SRGKN-IQGG 166 :: :: Db 124 KTSRVHLIVGVSPKIWEISDSINEGNNISTLCIATRPERETVMRIISPAAVFSESD 183 : Oy 167 KTLVSOLELDQDGWTCTVLQN-----OKVEFKDIVPPASALPAAPTSALEDDPOTA 221 :: :: Db 184 FYLERIGGITRRQSGETSCASNDAAPAVRKVNVTNYTPFIYS----EKGTGPVVGQG 239 : Oy 222 SALPDPPAASALPPA 236

```

Db          240  TLOC---ASAVPSA 251

RESULT 52
CONT_HUMAN
ID  _CONT_HUMAN  STANDARD;  .PRT;  1018 AA.
AC  Q12860; Q12861; Q14030;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Contactin precursor (Glycoprotein gp135).
GN  CNTN1.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_Taxid=9606;
RN  [1]
RP  SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC  TISSUE=Brain;
RX  MEDLINE=95048335; PubMed=7959734;
RA  Berglund E.O., Ranscht B.;
RT  "Molecular cloning and in situ localization of the human contactin
RL  gene (CNTN1) on chromosome 12q11-q12.";
RL  Genomics 21:571-582(1994).
[2]
RP  SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX  MEDLINE=94217459; PubMed=8164510;
RA  Reid R.A., Hemperly J.J.;
RT  "Identification and characterization of the human cell adhesion
RL  molecule contactin.";
RT  Brain Res. Mol. Brain Res. 21:1-8(1994).
CC  -1- FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS SYSTEM
CC  DEVELOPMENT. IN ASSOCIATION WITH CNTNAP1 SEEMS TO PLAY A ROLE IN
CC  THE FORMATION OF PARANODAL AXO-GLIAL JUNCTIONS IN MYELINATED
CC  PERIPHERAL NERVES AND MAY HAVE A ROLE IN THE SIGNALING BETWEEN
CC  AXONS AND MYELINATING GLIAL CELLS.
CC  -1- SUBUNIT: INTERACTS WITH CONTACTIN ASSOCIATED PROTEIN 1 IN CIS
CC  FORM. BINDS TO THE CARBONIC-ANHYDRASE LIKE DOMAIN OF PROTEIN-
CC  TYROSINE PHOSPHATASE ZETA (BY SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC  -1- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=2;
CC  Name=1;
CC  IsoId=Q12860-1; Sequence=Displayed;
CC  Name=2;
CC  IsoId=Q12860-2; Sequence=VSP_002500;
CC  -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC  -1- SIMILARITY: Contains 4 fibronectin type III domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration with
the Swiss Institute of Bioinformatics and the EMBL Outstation -
between the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/)
or send an email to license@isb-sib.ch.
-----
DR  EMBL; U07819; AAA67920.1; -
DR  EMBL; U07820; AAA67921.1; -
DR  EMBL; Z21488; CAA79696.1; -
DR  PIR; A54744; A54744.
DR  HSSP; P08921; 1A7B.
DR  Genew: HGNC:2171; CNTN1.
DR  MIM; 600016; -
DR  GO; GO:0005624; C:membrane fraction; TAS.
DR  InterPro; IPR008957; FN_III-like.
DR  InterPro; IPR003961; FN_III.
DR  InterPro; IPR007110; Ig-like.
DR  InterPro; IPR003598; Ig_C2.
DR  Pfam; PF00041; fn3; 4.
DR  Pfam; PF00047; Ig; 6.
DR  SMART; SM00060; FN3; 4.

```

DR	SMART; SMO0408; IGC2; 4.
KR	PROSITE; PS50835; IG LIKE; 6.
DW	Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
KV	Cell adhesion; Repeat; Alternative splicing; Polymorphism;
KM	Lipoprotein..
FT	SIGNAL 1 20 CONTRACTIN' REMOVED IN MATURE FORM.
FT	CHAIN 21 7993
FT	PROPEP 7994 1018 IG-LIKE C2-TYPE 1.
FT	DOMAIN 41 131 IG-LIKE C2-TYPE 2.
FT	DOMAIN 137 223 IG-LIKE C2-TYPE 3.
FT	DOMAIN 241 326 IG-LIKE C2-TYPE 4.
FT	DOMAIN 331 407 IG-LIKE C2-TYPE 5.
FT	DOMAIN 413 500 IG-LIKE C2-TYPE 6.
FT	DOMAIN 504 601 GLY/PRO-RICH.
FT	DOMAIN 602 609 FIBONECTIN TYPE-III 1.
FT	DOMAIN 609 710 FIBONECTIN TYPE-III 2.
FT	DOMAIN 711 812 FIBONECTIN TYPE-III 3.
FT	DOMAIN 813 908 FIBONECTIN TYPE-III 4.
FT	DOMAIN 909 1004 BY SIMILARITY.
FT	DISULFID 65 114 BY SIMILARITY.
FT	DISULFID 158 211 BY SIMILARITY.
FT	DISULFID 263 310 BY SIMILARITY.
FT	DISULFID 352 391 BY SIMILARITY.
FT	DISULFID 436 484 BY SIMILARITY.
FT	DISULFID 526 583 BY SIMILARITY.
FT	CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 437 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 473 473 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 494 494 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 521 521 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 933 933 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	LIPID 993 993 GPI-anchor amidated serine (Potential). Missing (in isoform 2).
FT	VANSPPLIC 21 31 /FTId=VSP_002500.
VT	Variant 798 798 V -> L (in dbSNP:1056020).
FT	/FTId=VAR_011722.
SEQ	SEQUENCE 1018 AA; 113320 MW; 4B8FDGCBPD43AED5 CRC64;
Query Match	Best Local Similarity 7.9%; Score 104; DB 1; Length 1018;
Matches 51; Conservative 27; Mismatches 90; Indels 82; Gaps 9;	Pred. No. 2.9;
Oy	16 LALLPAATQGS--NKKVLGRKGDGYELLTGAQSOKSIOGHMKNSNQIKLNGSGSLITGP 73
Db	409 LALAPTFEMNPWKXIIILAAGRGRIYECKPKAAPRPKYSW-----SKGT 452
Oy	74 SKLANDRADSRRLSDMGNFPLIILKLKLEDSDTYCEVEDEKEEYQLIVFGILTANSPTHL 133
Db	453 EMLVN--SSRLLIWMDGS--LEINNITTNDGGIYTICFAENNRKG-----ANS----- 495
Oy	134 LOGOSLVTLTLESP-----GSSPSVOGRS----- 157
Db	496 ----TGTLVIINDPTPIILLAPIADITVGEMNTMQCAASFDPALDLTFWSFNGVIDENK 551
Oy	158 -----PRKKNIGQGKITLSVSQLEDSTGWTCYVQNQKYVEFKDIYPASALPAPTG 212
Db	552 ENIHYNRMFLMSNGSELLIRNAQLHMRGRYTCTACTIVDNSSASADVLRGP--GPPPG 609
Oy	213 SALPDPOFAS 222
Db	610 LRIEDIRATS 619
RESULT 53	
ID CONT_MOUSE STANDARD; PRT; 1020 AA.	
AC PI2960;	
DT 01-JAN-1990 (Rel. 13; Created)	
DT 01-JAN-1990 (Rel. 13; Last sequence update)	
DT 10-OCT-2003 (Rel. 42; Last annotation update)	

DE Contactin precursor (Neural cell surface protein F3).
GN CNTN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Brain;
RC MEDLINE=89340657; PubMed=2475455;
RA Genattini G., Cibelli G., Rougon G., Mattei M.-G., Gordin C.;
RT "The mouse neuronal cell surface protein F3: a phosphatidylinositol-
anchored member of the immunoglobulin superfamily related to chicken
RT contactin.";
RT J. Cell Biol. 109:775-788(1989).
RL [2]
RN FUNCTION.
RP MEDLINE=21289248; PubMed=11395001;
RX Boyle M.E., Berzland E.O., Murai K.K., Weber L., Peles E., Ranscht B.;
RA "Contactin orchestrates assembly of the septate-like junctions at the
RT paranode in myelinated peripheral nerve.";
RT Neuron 30:385-397(2001).
CC -1- FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS SYSTEM
CC DEVELOPMENT. IN ASSOCIATION WITH CNTNAP1 SEEMS TO PLAY A ROLE IN
CC THE FORMATION OF PARANODAL AXO-GLIAL JUNCTIONS IN MYELINATED
CC PERIPHERAL NERVES AND MAY HAVE A ROLE IN THE SIGNALING BETWEEN
CC AXONS AND MYELINATING GLIAL CELLS.
CC -1- SUBUNIT: INTERACTS WITH CONTACTIN ASSOCIATED PROTEIN 1 IN CIS FORM
CC (BY SIMILARITY). Binds to the carbonic-anhydrase like domain of
CC protein-tyrosine phosphatase zeta (By similarity).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- MISCELLANEOUS: F3 SHARES WITH L1, N-CAM, MAG, AND OTHER CELL
CC ADHESION MOLECULES FROM NERVOUS TISSUE THE L2/HNK-1 CARBOHYDRATE
CC EPTOPE.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 4 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14943; CAA33075.1; -.
DR PIR: S05944; S05944.
DR HSSP: P08921; IABV.
DR MGD: MGI:105980; Cntn1.
DR InterPro: IPR008957; FN_III-like.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003598; IG_c2.
DR Pfam: PF00041; fn3; 4.
DR Pfam: PF00047; Ig; 6.
DR SMART: SMO0060; FN3; 4.
DR SMART: SMO0408; IGC2; 4.
DR PROSITE: PS50835; IG LIKE; 6.
DR Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
KM Cell adhesion; Repeat; Lipoprotein.
KW SIGNAL
FT 1
FT CHAIN 20
FT PROPEP 21 PROPEP 1020
FT DOMAIN 41 131
FT DOMAIN 137 223
FT DOMAIN 241 326
FT DOMAIN 313 407
FT DOMAIN 433 500
FT DOMAIN 504 603
FT DOMAIN 604 611
FT DOMAIN 611 712
FT DOMAIN 713 814
FT DOMAIN 815 910
CONTRACTIN.
REMOVED IN MATURE FORM.
IG-LIKE C2-TYPE 1.
IG-LIKE C2-TYPE 2.
IG-LIKE C2-TYPE 3.
IG-LIKE C2-TYPE 4.
IG-LIKE C2-TYPE 5.
IG-LIKE C2-TYPE 6.
GLY/PRO-RICH.
FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.
FIBRONECTIN TYPE-III 3.

FT	DOMAIN	911	1006	FIBRONECTIN TYPE-III 4.
FT	DISULFID	65	114	BY SIMILARITY.
FT	DISULFID	158	211	BY SIMILARITY.
FT	DISULFID	263	310	BY SIMILARITY.
FT	DISULFID	352	391	BY SIMILARITY.
FT	DISULFID	436	484	BY SIMILARITY.
FT	DISULFID	526	585	BY SIMILARITY.
FT	CARBOHYD	208	208	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	258	258	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	457	457	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	473	473	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	494	494	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	521	521	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	593	593	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	935	935	N-LINKED (GLCNAC. . .)
FT	LIPID	1001	1001	GPI-anchor amidated serine (Potential).
SO	SEQUENCE	1020 AA;	11338 MW;	9DCDA40EAA4CB67 CR64;
Query Match		7.9%;	Score 104;	DB 1;
Best Local Similarity		20.2%;	Pred. No. 2.9;	
Matches		51;	Conservative	28; Mismatches 89; Indels 84; Gaps 9
QY	16 LALLPAAAG--NKVVLGKGGDTVLTCTGASQKSIQPHMKNSNQIKLGNGSFLTGCP	73		
DB	409 LALATFEEMNPKKKILAAKGGRVITIECKPKAPPKFSW-----SKGT	452		
QY	74 SKLNRADSRSLQDGNFPFLITIKLIKIEDSPYICEVDOKEEYQLVFGLTANSDFHL	133		
DB	453 EMLVN--SSRIIWDGS--LEINNTNRDGGITCYCPAENNRGK-----ANS----	495		
QY	134 LGGOSLTLTLBSP-----GSSPSVQGRS-----	157		
DB	496 ----TGLTVITPTPTIILAPINADITVGENATMQCAASFPDALDTFWMSFNGYVIDENK	551		
QY	158 -----PRGNIOGGKTLVSQGLQDSGTTCTVNLQNKVKERKIDIVPASALPAP	210		
DB	552 EITHIHVGRNFMLDANGELLIRNAQLKGAGRYTCTAQTITVDSASADLVVRGP--FGPP	609		
QY	211 TGSALPDPTAS	222		
DB	610 GGLREDIRATS	621		

RESULT 54

CONT RAT

ID	CONT	RAT	STANDARD;	PRT;	1021 AA.
AC	O63198;				
DC	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Contactin precursor (Neural adhesion molecule F3).				
GN	CNTN1.				
OS	Rattus norvegicus (rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_Taxid=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Miscar; TISSUE=Brain;				
RC	MEDLINE=95295987; PubMed=7777204;				
RA	Hosoya H., Shimazaki K., Kobayashi S., Takahashi H., Shirasawa T.,				
RA	Takenawa T., Watanabe K.;				
RT	"Developmental expression of the neural adhesion molecule F3 in the				
RT	rat brain."				
RL	Neurosci. Lett. 186:83-86(1995).				
RN	[2]				
RP	INTERACTION WITH PTPRZ1.				
RC	MEDLINE=95354206; PubMed=7628014;				
RA	Petes E., Nativ M., Campbell P.L., Sakurai T., Martinez R., Lev S.,				
RA	Clary D.O., Shilling J., Barnea G., Plowman G.D., Grumet M.,				
RT	Schlesinger J.;				
RT	"The carboxin anhydride domain of receptor tyrosine phosphatase beta				

RT		is a functional ligand for the axonal cell recognition molecule
RT		contactin.";
RL		Cell 82:251-260(1995).
RN	[3]	
RP		INTERACTION WITH CNTNAP1.
RX		MEDLINE=97224125; PubMed=9118959;
RY		Peles E., Nativ M., Luetzig M., Grunnet M., Schilling J., Martinez R.,
RA		Ploewman G.D., Schlesinger J.;
RT		"Identification of a novel contactin-associated transmembrane receptor
RT		with multiple domains implicated in protein-protein interactions.";
RL		JMBIO J. 16:978-988(1997).
CC	-I-	FUNCTION: MEDIATES CELL SURFACE INTERACTIONS DURING NERVOUS SYSTEM
CC		DEVELOPMENT. IN ASSOCIATION WITH CNTNAP1 SEEMS TO PLAY A ROLE IN
CC		THE FORMATION OF PARANODAL AYO-GLIAL JUNCTIONS IN MYELINATED
CC		PERIPHERAL NERVES AND MAY HAVE A ROLE IN THE SIGNALING BETWEEN
CC		AXONS AND MYELINATING GLIAL CELLS.
CC	-I-	SUBUNIT: INTERACTS WITH CONTACTIN ASSOCIATED PROTEIN 1 IN CIS
CC		FORM. BINDS TO THE CARBONIC-ANHYDRASE LIKE DOMAIN OF PROTEIN-
CC		TYROSINE PHOSPHATASE ZETA.
CC	-I-	SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC	-I-	SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC	-I-	SIMILARITY: Contains 4 fibronectin type III domains.
CC		
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outcation -
CC		the European Bioinformatics institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See http://www.ibb-sib.ch/announce/
CC		or send an email to license@ibb-sib.ch).
CC		-----
DR	EMBL:	D38492; BAA07504.1; .
DR	HSSP:	P40189; 18QU.
DR	InterPro:	IIPR008957; FN_III-like.
DR	InterPro:	IIPR003961; FN_III.
DR	InterPro:	IIPR007110; IG_II-like.
DR	InterPro:	IIPR003598; IG_C2.
DR	Pfam:	PF00047; fn3; 4.
DR	Pfam:	PF00047; ig; 6.
DR	SMART:	SMO0060; FN3; 4.
DR	SMART:	SMO0408; IGC2; 4.
DR	PROSITE:	PS50835; IG_LIKE; 6.
KW		Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
KM		Cell adhesion; Repeat; Lipoprotein.
FT	SIGNAL	1 POTENTIAL.
FT	CHAIN	21 ?1001 CONTRACTIN.
FT	PROPEP	?1002 1021 REMOVED IN MATURE FORM (POTENTIAL).
FT	DOMAIN	41 131 IG-LIKE C2-TYPE 1.
FT	DOMAIN	137 223 IG-LIKE C2-TYPE 2.
FT	DOMAIN	241 326 IG-LIKE C2-TYPE 3.
FT	DOMAIN	331 407 IG-LIKE C2-TYPE 4.
FT	DOMAIN	413 500 IG-LIKE C2-TYPE 5.
FT	DOMAIN	504 603 IG-LIKE C2-TYPE 6.
FT	DOMAIN	604 611 GLV/PRO-RICH.
FT	DOMAIN	611 712 FIBRONECTIN TYPE-III 1.
FT	DOMAIN	713 814 FIBRONECTIN TYPE-III 2.
FT	DOMAIN	815 910 FIBRONECTIN TYPE-III 3.
FT	DOMAIN	911 1006 FIBRONECTIN TYPE-III 4.
FT	DISULFID	65 114 BY SIMILARITY.
FT	DISULFID	158 211 BY SIMILARITY.
FT	DISULFID	263 310 BY SIMILARITY.
FT	DISULFID	352 391 BY SIMILARITY.
FT	DISULFID	436 484 BY SIMILARITY.
FT	DISULFID	526 585 BY SIMILARITY.
FT	CARBOHYD	208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	457 457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	473 473 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	494 494 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	521 521 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	593 593 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	935 935 N-LINKED (GLCNAC. .) (POTENTIAL).

FT LIPID 1001 1001 GPI-anchor amidated serine (Potential).
SQ SEQUENCE 1021 AA; 113494 MW; FCDCL1055EE5C68 CRC64;
Query Match 7.9%; Score 104; DB 1; Length 1021;
Best Local Similarity 20.2%; Pred. No. 3;
Matches 51; Conservative 28; Mismatches 89; Indels 84; Gaps 9;
QY 16 LALLPAATQG--NKVVLGKGGDYVELTCTASQKSIQFHWKNSNQIKLNGQSFLTKGP 73
DB 409 LALAPFEEMPMKKKILAAAGKGVILECKEAPKPEFSW-----SKGT 452
QY 74 SKLNDRADSRSLMDQNPFLIKNKLEIDSDTYICEVEQKEVQLVGLTANSPDTH 133
DB 453 EMLVN--SSRIILWEDGS--LEINNITRNDGCIYTCFAENNRK-----ANS----- 495
QY 134 LOGQSLLTLESPP-----GSSPSVQCRS----- 157
DB 496 ---TGLVITNPFRILAPINDITVGENATMQCASFPDPSLDLFFVMSFNGVIDFNK 551
QY 158 -----PRGNIOGGKTLVSQLELDSDGTWCTVQONQKVEFKIDIVPRASALPAPP 210
DB 552 EITNHYQRFEMDANGELLIRNAQLKAGRYCTACTIVDNSSASADLVVRGP--PGPP 609
QY 211 TGSALPDPOFAS 222
DB 610 GGLRIEDIRATS 621
RESULT 55
KVSS MOUSE STANDARD; PRT; 108 AA.
AC P01652;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region J606.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Slankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
binding proteins."
RL J. Immunol. 128:302-307(1982).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR; A92811; KWS06.
DR HSSP; P01607; IREI.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 FRAMEWORK-2.
FT DOMAIN 35 49 FRAMEWORK-3.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON TER 108
SQ SEQUENCE 108 AA; 11810 MW; BDE4DD31076F2AFB CRC64;
Query Match 7.9%; Score 103.5; DB 1; Length 108;
Best Local Similarity 39.0%; Pred. No. 0.2;
Matches 30; Conservative 5; Mismatches 37; Indels 5; Gaps 2;
QY 34 GDVETLCTASQKSIQFHWKNSNQIKLNGQSFLTKGPSKLNDRADSRSLMDQ--NF 92

DB 16 GDIYVMTQASGTSINLWFOQKP-----GKAPKLLIYASNLDEGVPSRFGSGRYCTGF 71
QY 93 PLIRNKLEIDSDTYIC 109
DB 72 TLTIISLEDEDMATYFC 88
RESULT 56
KVW HUMAN STANDARD; PRT; 129 AA.
ID K1VW HUMAN
AC P04431;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Walker precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85014148; PubMed=6091049;
RA Klobeck H.G., Combrato G., Zachau H.G.;
RT "Immunoglobulin genes of the kappa light chain type from two human
RT lymphoid cell lines are closely related."
RL Nucleic Acids Res. 12:6995-7006(1984).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X00965; CAA25477.1; ALY_TERM.
DR PIR; A01883; K1HWK.
DR HSSP; P01607; IREI.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0003823; P:antigen binding; NAS.
DR GO; GO:0006955; P:immune response; NAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 22
FT CHAIN 23 129 IG KAPPA CHAIN V-I REGION WALKER.
FT DOMAIN 23 45 FRAMEWORK-1.
FT DOMAIN 46 56 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 57 71 FRAMEWORK-2.
FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 79 110 FRAMEWORK-3.
FT DOMAIN 111 119 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 120 128 FRAMEWORK-4.
FT DISULFID 45 110 BY SIMILARITY.
FT NON TER 129
SQ SEQUENCE 129 AA; 14069 MW; F941FA07DAFC2F9 CRC64;
Query Match 7.9%; Score 103.5; DB 1; Length 129;
Best Local Similarity 32.8%; Pred. No. 0.25; Indels 17; Gaps 5;
Matches 43; Conservative 10; Mismatches 61;
QY 1 MNRGVPFRHLVLTQALPAA-----TQGNKVVLGKGGDYVELTCTASQKSIQFHWK 54
DB 1 MNRKVRPAQ--LLGLLLWLRGARCIDIQMTQSPESLSASVDRVTITCRASQSIYNLWY 58
QY 55 NSNQIKLNGQSFLTKGPSKLNDRADSRSLMDQ--NPLIINKLEIDSDTYICEVED 113
DB 59 QQKP-----GKAPKLLIYASNLQSGVTSRFGSGGSDTFLTITSSLPQEDSATYTC----- 110

```

Oy 114 QKEVQLLVFG 124
Db 111 QOSYSLITLTFG 121

RESULT 57
CD7_MOUSE STANDARD; PRT; 210 AA.
AC P50283;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE T-cell antigen CD7 precursor.
GN CD7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91138718; PubMed=7678579;
RA Yoshikawa K., Seto M., Ueda R., Obata Y., Fukatsu H., Segawa A.,
RT Takahashi T.;
RL "Isolation and characterization of mouse CD7 cDNA.";
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=129;
RC MEDLINE=94164701; PubMed=7509775;
RA Lee D.M., Watson M.L., Seldin M.F.;
RT "Mouse Cd7 maps to chromosome 11.";
RL Immunogenetics 39:289-290(1994).
RN [3]
RP SEQUENCE OF 1-24 AND 176-210 FROM N.A.
RC STRAIN=BALB/c; TISSUE=Spleen;
RA MEDLINE=95104926; PubMed=7528728;
RA Yoshikawa K., Seto M., Ueda R., Obata Y., Aoki S., Takahashi T.;
RT "Molecular cloning of the gene coding for the mouse T-cell antigen
RT CD7.";
RL Immunogenetics 41:159-161(1995).
RN [4]
RP INTERACTION WITH SECTM1.
RX MEDLINE=20119303; PubMed=10652336;
RA Lyman S.D., Escobar S., Rousseau A.-M., Armstrong A., Fanslow W.C.;
RT "Identification of CD7 as a cognate of the human K12 (SECTM1)
RT protein.";
RL J. Biol. Chem. 275:3431-3437(2000).
CC -1- FUNCTION: Not yet known.
CC -1- SUBUNIT: Interacts with SECTM1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; D10329; BAA01171.1; -
DR EMBL; U23462; AAB17482.1; -
DR EMBL; D31956; BAA06728.1; -
DR EMBL; D31957; BAA06728.1; JOINED.
DR EMBL; D31958; BAA06728.1; JOINED.
DR EMBL; D31959; BAA06728.1; JOINED.
DR PIR; I49294; I49294.
DR HSSP; P80362; 1WTL.
DR MCD; MGI:88344; Cd7.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.

```

```

DR PROSITE; PS50835; IG LIKE; 1.
KW T-cell; Signal; Immune response; Antigen; Transmembrane; Glycoprotein;
KW Immunoglobulin domain; Receptor; Lipoprotein; Palmitate.
FT SIGNAL 1 23
FT CHAIN 24 210
FT DOMAIN 24 150 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 151 171 POTENTIAL.
FT DOMAIN 172 210 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 24 129 IG-LIKE.
FT DISULFD 45 111 POTENTIAL.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 168 168 S-palmitoyl cysteine (By similarity).
FT CONFLICT 69 69 P = L (IN REF. 2).
SQ SEQUENCE 210 AA; 23152 MW; 0EBD968B302962B CRC64;

Query Match 7.9%; Score 103.5; DB 1; Length 210;
Best Local Similarity 21.7%; Pred. No. 0.46;
Matches 57; Conservative 32; Mismatches 61; Indels 113; Gaps 13;

Oy 10 LLLVQLA-LIPA-----ATQGNKVVLGKGDVVELCTTASQKKSITQFMKNSNQIKIL 62
Db 7 LALLLTLAGILPGLDADVDVHQSPLTASGDSVNITCSYR----- 48
Oy 63 GNGSGFLTKGSPKLNDRADSRSLMDQGNFLLIKNLIKEDSVTYICEVEDQKEEVQLLV 122
Db 49 GHLGILMK-----KIWPA-----YNYVFEKRE----- 74
Oy 123 FGLTRANSDTHLQGSULTLTPLESPGSSPSVQCRSPRGK-NIQGK---TLVSQLELD 178
Db 75 -----PTVD-RTFGRINFSQSKNLTITISLQAD 105
Oy 179 SGWTCVVLQNKVY---FKIDIVPRASALPAPPGSALPDQTSALPDPPAAS-AL 233
Db 106 TGDYTCBAV---RKVSARGLFTTVVKEKSS-----QEAYRSQEPLOTSSSF 149
Oy 234 PALAVSFLGLGIGVAVCLAR 256
Db 150 PAATAVGFFFTGLLGVCSMLR 172

RESULT 58
NRG_DROME STANDARD; PRT; 1302 AA.
ID NRG_DROME
AC P20241; O61541; O61542; Q24414; Q24415; Q95064; Q9V3X0;
DT 01-FEB-1991 (Rel. 17, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neuroglial precursor.
GN NRG OR CGI634.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SHORT), AND SEQUENCE OF 24-41 AND 737-751.
RX MEDLINE=90030418; PubMed=2805067;
RA Bieber A.J., Snow P.M., Hortsch M., Patel N.H., Jacobs J.R.,
RA Treguina Z.R., Schilling J., Goodman C.S.;
RT "Drosophila neuroglial: a member of the immunoglobulin superfamily
RT with extensive homology to the vertebrate neural adhesion molecule
RT L1.";
RL Cell 59:447-460(1989).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98332718; PubMed=9666073;
RA Zhao G., Hortsch M.;
RT "The analysis of genomic structures in the L1 family of cell adhesion
RT proteins provides no evidence for exon shuffling events after the
RT separation of arthropod and chordate lineages.";
RL Gene 215:47-55(1998).

```


[3]
 RP REVISIONS.
 RA Hortsch M.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amentides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Baau A.V., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Baerentzen K.Y., Benson P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Delcher A.S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunlov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fowler C., Gabriellian A.E., Gacy N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heilmann T.J., Hernandez J.R., Houck J.,
 RA Houston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao X., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Moparthy C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman K.A., Pan S., Pollard J., Porri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Scapleton M., Strong R., Sun E.,
 RA Svaykbas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2185-2195(2000).
 [5]
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC STRAIN=Berkley; TISSUE=leadi;
 RA Scapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuccio J., Pacle J., Paragass V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RL submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 [6]
 RP SEQUENCE OF 1182-1302 FROM N.A., FUNCTION, ALTERNATIVE SPLICING, AND
 RP TISSUE SPECIFICITY.
 RC TISSUE=Embryo;
 RX MEDLINE=90262720; PubMed=1693086;
 RA Hortsch M., Bieber A.J., Patel N.H., Goodman C.S.;
 RT "Differential splicing generates a nervous system-specific form of
 RA *Drosophila neuroglian*.";
 RL Neuron 4:697-703(1990).
 [7]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 610-814.
 RX MEDLINE=94213741; PubMed=1512815;
 RA Huber A.H., Wang Y.-M.E., Bieber A.J., Bjorkman P.J.;
 RT "Crystal structure of tandem type III fibronectin domains from
 RA *Drosophila neuroglian* at 2.0 Å";
 RL Neuron 12:717-731(1994).
 [8]
 RP -1- FUNCTION: The long isoform may play a role in neural and glial
 RP cell adhesion in the developing embryo. The short isoform may be a

[illegible]

FT	CARBOHYD	652	652		N-LINKED (GLCNAC . . .) (POTENTIAL) .
FT	CARBOHYD	683	683		N-LINKED (GLCNAC . . .) (POTENTIAL) .
FT	CARBOHYD	821	821		N-LINKED (GLCNAC . . .) (POTENTIAL) .
FT	CARBOHYD	1125	1125		N-LINKED (GLCNAC . . .) (POTENTIAL) .
FT	VARSPLIC	1224	1239		OPTEDGSPFGQYVPGK -> MNEDGSPFGQYGRKG. (In isoform short)
FT					/FTID=VSP_002601.
FT	VARSPLIC	1240	1302		Missing (in isoform short) .
FT					/FTID=VSP_002602.
FT	CONFLICT	85	86		NR -> KP (IN REF. 2) .
FT	CONFLICT	1282	1282		MISSING (IN REF. 6) .
FT	STRAND	619	625		
FT	STRAND	629	635		
FT	TURN	640	641		
FT	STRAND	646	653		
FT	TURN	657	658		
FT	STRAND	661	668		
FT	TURN	669	670		
FT	STRAND	673	677		
FT	STRAND	682	692		
FT	TURN	693	694		
FT	STRAND	695	696		
FT	STRAND	706	708		
FT	STRAND	721	723		
FT	TURN	727	728		

Query Match Best Local Similarity 20.7%; Pred. No. 4.3;
Matches 54; Conservative 38; Mismatches 104; Indels 65; Gaps 13

Qy	25	GKRVVLGGKKDVTVELTCTASQCK-SIQPHMKNISQIKLNGQS-FLTKGPSKLNDRA-80
Db	225	GKRVLL-----DVCKMGVASQMKHPHVQYRSRSLALRGKRMELFCITGGTPPPQTW-280
Qy	81	--DSRRSLMD-----QGNF--PLIINKLKIEDSDTYICEVEDQKEEVOVLVEGLTANSPTH-132
Db	281	SKDGGRIGMSDRITGGHGKSLVINTQTFDDGITTCVDINSNGNAQSPSIIILNVSVPY-340
Qy	133	LLOGSLTLTTESPPGSSPSVQCGRS-----PRGNKIQCGK-----TLSV-171
Db	341	FTKEPELTATMED---EEVVPECRAAGVPEPKISMTHNGKPLEQSTPNRPRTVTDNTIRI-397
Qy	172	SQLELDQSGTWCTVLQNQKVEFKIDIVPRASALPAPTGSALPDQTASALDPPEAAS-231
Db	398	INLVGDGTGVNGCNA-TNSLGTVVK-DVYLNVQA-----EPPTIS-435
Qy	232	ALPALAVISFLLGLGLGVAC-252
Db	436	EAPAAVSTVD--GRNVTIKC-453

RESULT 59
ECCTO_RAT ECCTO_RAT STANDARD; PRT; 519 AA.
AC P16573;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Eccto-AtPase precursor (Cell-CAM 105) (C-CAM 105) (ATP-dependent
DE laurocolate-carrier protein) (GP110).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_Taxid=10116;
[1]
SEQUENCE FROM N.A. (ISOFORM LONG), AND SEQUENCE OF 50-68.
RN RP STRAIN=Sprague-Dawley; TISSUE=Liver;
RC MEDLINE=893140561; PubMed=2527235;
RA Lin S.-H., Guidotti G.;
RT "Cloning and expression of a cDNA coding for a rat liver plasma
membrane eccto-ATPase. The primary structure of the eccto-ATPase is
similar to that of the human biliary glycoprotein I.",
J. Biol. Chem. 264:14408-14414(1989).

RN RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC STRAIN=Sprague-Dawley, and wistar; TISSUE=Liver;
 RX MEDLINE=93279310; PubMed=8504806;
 RA Edlund M., Gaardsvoll H., Bock E., Oebrigk B.;
 RT "Different isoforms and stock-specific variants of the cell adhesion
 molecule C-CAM (cell-CAM 105) in rat liver.";
 RL Eur. J. Biochem. 213:1109-1116(1993).
 RN RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=92344597; PubMed=1637321.
 RA Culic O., Huang Q., Flanagan D., Hixon D., Lin S.-H.;
 RT "Molecular cloning and expression of a new rat liver cell-CAM105
 isoform. Differential phosphorylation of isoforms.";
 RL Biochem. J. 285:47-53(1992).
 RN RP [4]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Intestine;
 RX MEDLINE=94089980; PubMed=8240240;
 RA Cheung P.H., Culic O., Qiu Y., Earley K., Thompson N., Hixson D.C.,
 RA Lin S.-H.;
 RT "The cytoplasmic domain of C-CAM is required for C-CAM-mediated
 adhesion function: studies of a C-CAM transcript containing an
 unspliced intron.";
 RL Biochem. J. 295:427-435(1993).
 RN RP [5]
 RC PARTIAL SEQUENCE.
 RP MEDLINE=90292222; PubMed=2141577;
 RX Aurtivillius M., Hansen O.C., Lazrek M.B.S., Bock E., Oebrigk B.;
 RA "The cell adhesion molecule Cell-CAM 105 is an ecto-ATPase and a
 member of the immunoglobulin superfamily.";
 RT FEBS Lett. 264:267-269(1990).
 RN RP [6]
 RC PARTIAL SEQUENCE.
 RP STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=93292517; PubMed=8513803;
 RA Becker A., Lucka L., Kilian C., Kannicht C., Reutter W.;
 RT "Characterisation of the ATP-dependent taurocholate-carrier protein
 (gp110) of the hepatocyte canalicular membrane.";
 RL Eur. J. Biochem. 214:539-548(1993).
 RN RP [7]
 RC ALTERNATIVE SPLICING.
 RP MEDLINE=93123234; PubMed=8380406;
 RX Najjar S.M., Accilli D., Philippe N., Jernberg J., Margolis R.,
 RA Taylor S.I.;
 RT "gp120/ecto-ATPase, an endogenous substrate of the insulin receptor
 tyrosine kinase, is expressed as two variably spliced isoforms.";
 RL J. Biol. Chem. 268:1201-1206(1993).
 RN RP [8]
 RC CHARACTERIZATION.
 RP MEDLINE=91354197; PubMed=1831973;
 RA Lin S.-H., Culic O., Flanagan D., Hixson D.C.;
 RT "Immunohistochemical characterization of two isoforms of rat liver ecto-
 ATPase that show an immunological and structural identity with a
 glycoprotein cell-adhesion molecule with Mr 105,000.";
 RL Biochem. J. 278:155-161(1991).
 RN RP [9]
 RC FUNCTION. THIS PROTEIN IS A CALCIUM-INDEPENDENT CELL ADHESION
 MOLECULE WITH HOMOPHILIC BINDING PROPERTIES. MAY PLAY A ROLE IN
 THE FORMATION AND MAINTENANCE OF THE SPECIALIZED MEMBRANE
 STRUCTURE OF THE APICAL SURFACE OF THE HEPATOCYTES.
 CC [10]
 CC [11]
 CC [12]
 CC [13]
 CC [14]
 CC [15]
 CC [16]
 CC [17]
 CC [18]
 CC [19]
 CC [20]
 CC [21]
 CC [22]
 CC [23]
 CC [24]
 CC [25]
 CC [26]
 CC [27]
 CC [28]
 CC [29]
 CC [30]
 CC [31]
 CC [32]
 CC [33]
 CC [34]
 CC [35]
 CC [36]
 CC [37]
 CC [38]
 CC [39]
 CC [40]
 CC [41]
 CC [42]
 CC [43]
 CC [44]
 CC [45]
 CC [46]
 CC [47]
 CC [48]
 CC [49]
 CC [50]
 CC [51]
 CC [52]
 CC [53]
 CC [54]
 CC [55]
 CC [56]
 CC [57]
 CC [58]
 CC [59]
 CC [60]
 CC [61]
 CC [62]
 CC [63]
 CC [64]
 CC [65]
 CC [66]
 CC [67]
 CC [68]
 CC [69]
 CC [70]
 CC [71]
 CC [72]
 CC [73]
 CC [74]
 CC [75]
 CC [76]
 CC [77]
 CC [78]
 CC [79]
 CC [80]
 CC [81]
 CC [82]
 CC [83]
 CC [84]
 CC [85]
 CC [86]
 CC [87]
 CC [88]
 CC [89]
 CC [90]
 CC [91]
 CC [92]
 CC [93]
 CC [94]
 CC [95]
 CC [96]
 CC [97]
 CC [98]
 CC [99]
 CC [100]
 CC [101]
 CC [102]
 CC [103]
 CC [104]
 CC [105]
 CC [106]
 CC [107]
 CC [108]
 CC [109]
 CC [110]
 CC [111]
 CC [112]
 CC [113]
 CC [114]
 CC [115]
 CC [116]
 CC [117]
 CC [118]
 CC [119]
 CC [120]
 CC [121]
 CC [122]
 CC [123]
 CC [124]
 CC [125]
 CC [126]
 CC [127]
 CC [128]
 CC [129]
 CC [130]
 CC [131]
 CC [132]
 CC [133]
 CC [134]
 CC [135]
 CC [136]
 CC [137]
 CC [138]
 CC [139]
 CC [140]
 CC [141]
 CC [142]
 CC [143]
 CC [144]
 CC [145]
 CC [146]
 CC [147]
 CC [148]
 CC [149]
 CC [150]
 CC [151]
 CC [152]
 CC [153]
 CC [154]
 CC [155]
 CC [156]
 CC [157]
 CC [158]
 CC [159]
 CC [160]
 CC [161]
 CC [162]
 CC [163]
 CC [164]
 CC [165]
 CC [166]
 CC [167]
 CC [168]
 CC [169]
 CC [170]
 CC [171]
 CC [172]
 CC [173]
 CC [174]
 CC [175]
 CC [176]
 CC [177]
 CC [178]
 CC [179]
 CC [180]
 CC [181]
 CC [182]
 CC [183]
 CC [184]
 CC [185]
 CC [186]
 CC [187]
 CC [188]
 CC [189]
 CC [190]
 CC [191]
 CC [192]
 CC [193]
 CC [194]
 CC [195]
 CC [196]
 CC [197]
 CC [198]
 CC [199]
 CC [200]
 CC [201]
 CC [202]
 CC [203]
 CC [204]
 CC [205]
 CC [206]
 CC [207]
 CC [208]
 CC [209]
 CC [210]
 CC [211]
 CC [212]
 CC [213]
 CC [214]
 CC [215]
 CC [216]
 CC [217]
 CC [218]
 CC [219]
 CC [220]
 CC [221]
 CC [222]
 CC [223]
 CC [224]
 CC [225]
 CC [226]
 CC [227]
 CC [228]
 CC [229]
 CC [230]
 CC [231]
 CC [232]
 CC [233]
 CC [234]
 CC [235]
 CC [236]
 CC [237]
 CC [238]
 CC [239]
 CC [240]
 CC [241]
 CC [242]
 CC [243]
 CC [244]
 CC [245]
 CC [246]
 CC [247]
 CC [248]
 CC [249]
 CC [250]
 CC [251]
 CC [252]
 CC [253]
 CC [254]
 CC [255]
 CC [256]
 CC [257]
 CC [258]
 CC [259]
 CC [260]
 CC [261]
 CC [262]
 CC [263]
 CC [264]
 CC [265]
 CC [266]
 CC [267]
 CC [268]
 CC [269]
 CC [270]
 CC [271]
 CC [272]
 CC [273]
 CC [274]
 CC [275]
 CC [276]
 CC [277]
 CC [278]
 CC [279]
 CC [280]
 CC [281]
 CC [282]
 CC [283]
 CC [284]
 CC [285]
 CC [286]
 CC [287]
 CC [288]
 CC [289]
 CC [290]
 CC [291]
 CC [292]
 CC [293]
 CC [294]
 CC [295]
 CC [296]
 CC [297]
 CC [298]
 CC [299]
 CC [300]
 CC [301]
 CC [302]
 CC [303]
 CC [304]
 CC [

CC	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	
----	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

[illegible]

DR Pfam; PF00047; Ig; 2.
 DR SMART; SMO0409; IG; 3.
 DR PROSITE; PSS0835; IG-LIKE; 2.
 KM Signal; Transmembrane; Glycoprotein; Immunoglobulin domain; Repeat.
 FT SIGNAL 1 22
 FT CHAIN 23 521
 FT DOMAIN 23 442
 FT TRANSMEM 443 463
 FT DOMAIN 464 521
 FT DOMAIN 37 163
 FT DOMAIN 165 246
 FT DOMAIN 258 341
 FT DOMAIN 345 412
 FT DOMAIN 505 518
 FT DISULFID 44 156
 FT DISULFID 185 235
 FT DISULFID 276 327
 FT DISULFID 363 405
 FT CARBOHYD 184 184
 FT CARBOHYD 244 244
 FT CARBOHYD 309 309
 FT CARBOHYD 337 337
 FT CARBOHYD 381 381
 SQ SEQUENCE 521 AA; 56977 MM; 64C29C40EB780E48 CRC64;
 Query Match 7.8%; Score 103; DB 1; Length 521;
 Best Local Similarity 24.9%; Pred. No. 1.5; Mismatches 96; Indels 32; Gaps 11;
 Matches 55; Conservative 38; Mismatches 96; Indels 32; Gaps 11;
 QY 5 VPFRHLVLQALALPAATQGNKVLGKGGDTVELTCTASQ-KKSIGPFM-KNSNQIKIL 62
 DB 149 LFNALSGSLRLRGQASMINASPSGVLTLSQWVLNCSFSPRDPVSHWVGQGNRPVY 208
 QY 63 GNQGSFLTKGSKLNDRAISRSLMDQGNFPLIK-----NLKEDSTYICEVEDQKEE 117
 DB 209 NSPHFLAETFLLL-----POVSPIDSGTCGCVLTFRDGFVSI-----TYNLKVLGLEPV 259
 QY 118 VQLVFE---GLTANSDTLLG---QSLTLTLESPPSSPVQCRSPKXKIOGKTLISV 171
 DB 260 APLTVVAEGSRVELPCHLPVGVTPSLIKMKPVGPGGFLPV---AGKS--GNFTLHL 314
 QY 172 SOLELDGSGTWTCTV-LQNRK---KYEFKIDIVPRASALP 207
 DB 315 EAVGLAQAGTTCGSIHLOGQGLNMTVTLAVTVTPKSGLP 355
 RESULT 61
 CD22_MOUSE STANDARD: PRT: 862 AA.
 ID CD22_MOUSE
 AC P35329; O9JUH2; O9JUX9; O9JUY0; O9JUY1; Q9R056; Q9R094; Q9WU51;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE B-cell receptor CD22 precursor (Ieu-14) (B-lymphocyte cell adhesion molecule) (BR-CAM) (Siglec-2).
 GN CD22 OR LYB-8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_Taxid=10090;
 RX NCB1_Taxid=10090;
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2J, and BALB/c; TISSUE=Liver;
 RX MEDLINE=93315834; PubMed=8100843;
 RA Brenan C.I., Copeland R.M., Sundberg H.A., Parkhouse R.M.,
 RA Brenan C.I., Copeland N.G., Jenkins N.A., Clark E.A.,
 RT "Organization of the murine Cd22 locus. Mapping to chromosome 7 and
 RT characterization of two alleles."
 RL J. Immunol. 151:175-187(1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6, BxSB, and MRL;
 RX MEDLINE=99432012; PubMed=10501843;

RA Lajunias F., Ikon-Zekri N., Fossati Jimack L., Chicheportiche Y.,
 RA Parkhouse R.M., Mary C., Reininger L., Britchouse G., Izui S.,
 RT "Polymorphisms in the Cd22 gene of inbred mouse strains."
 RL Immunogenetics 49:991-995(1999).
 RN [3]
 RP SEQUENCE OF 1-178 FROM N.A. (ISOFORMS 1; 2 AND 3).
 RC STRAIN=NZW; TISSUE=Spleen;
 RX MEDLINE=20432275; PubMed=10975807;
 RA Mary C., Laporte C., Parzy D., Santiago M.L., Stefani F.,
 RA Lajunias F., Parkhouse M.E., O'Keefe T.L., Neuberger M.S., Izui S.,
 RA Reininger L.,
 RT "Dysregulated expression of the Cd22 gene as a result of a short
 RT interspersed nucleotide element insertion in Cd22alpha lupus-prone
 RT mice."
 RL J. Immunol. 165:2987-2996(2000).
 RN [4]
 RP SIALIC ACID BINDING.
 RX MEDLINE=95179521; PubMed=7533044;
 RA Keim S., Pelz A., Schauer R., Filbin M.T., Tang S., de Bellard M.E.,
 RA Schnaar R.L., Mahoney J.A., Hartnell A., Bradfield P., Crocker P.R.,
 RT "Sialoadhesin, myelin-associated glycoprotein and CD22 define a new
 RT family of sialic acid-dependent adhesion molecules of the
 RT immunoglobulin superfamily."
 RL Curr. Biol. 4:965-972(1994).
 RN [5]
 RP INTERACTION WITH GRB2; SYK; PIK3R1/PIK3R2 AND PLCG1, PHOSPHORYLATION
 RP OF TYR-777, TYR-822, TYR-837 AND TYR-857, AND MUTAGENESIS OF TYR-822.
 RX MEDLINE=99303650; PubMed=10373493;
 RA Yohannan J., Wienands J., Coggeshall K.M., Justement L.B.,
 RT "Analysis of tyrosine phosphorylation-dependent interactions between
 RT stimulatory effector proteins and the B cell co-receptor CD22."
 RL J. Biol. Chem. 274:18769-18776(1999).
 RN [6]
 RP INTERACTION WITH GRB2, SHC1 AND INPP5D, AND PHOSPHORYLATION BY LYN.
 RX MEDLINE=20298808; PubMed=10748054;
 RA Poe J.C., Fujimoto M., Jansen P.J., Miller A.S., Tedder T.F.,
 RT "CD22 forms a quaternary complex with SHP, Grb2, and Shc. A pathway
 RT for regulation of B lymphocyte antigen receptor-induced calcium
 RT flux."
 RL J. Biol. Chem. 275:17420-17427(2000).
 CC -1- FUNCTION: Mediates B-cell B-cell interactions. May be involved in
 CC the localization of B-cells in lymphoid tissues. Binds sialylated
 CC glycoproteins; one of which is CD45. Preferentially binds to
 CC alpha2,6-linked sialic acid. The sialic acid recognition site can
 CC be masked by cis interactions with sialic acids on the same cell
 CC surface. Upon ligand induced tyrosine phosphorylation in the
 CC immune response seems to be involved in regulation of B cell
 CC antigen receptor signaling. Plays a role in positive regulation
 CC through interaction with Src family tyrosine kinases and may also
 CC act as an inhibitory receptor by recruiting cytoplasmic
 CC phosphatases via their SH2 domains that block signal transduction
 CC through dephosphorylation of signaling molecules.
 CC -1- SUBUNIT: Interacts with SYK, PIK3R1/PIK3R2, PLCG1, SHC1, INPP5D
 CC and GRB2 upon phosphorylation. May form a complex with
 CC INPP5D/SHP, GRB2 and SHC1. Interacts with PTPN6/SHP-1 upon
 CC phosphorylation (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1; Synonyms=CD22-beta;
 CC IsoId=P35329-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P35329-2; Sequence=VSP_002532;
 CC Name=3;
 CC IsoId=P35329-3; Sequence=VSP_002533;
 CC -1- TISSUE SPECIFICITY: B lymphocytes.
 CC -1- DOMAIN: Contains 3 copies of a cytoplasmic motif that is referred
 CC to as the immunoreceptor tyrosine-phased inhibitor motif (ITIM).
 CC This motif is involved in downmodulation of cellular responses.
 CC The phosphorylated ITIM motif binds to the SH2 domain of
 CC PTPN6/SHP-1.
 CC -1- PTM: Phosphorylated on tyrosine residues by LYN (Probable).
 CC -1- PTM: Phosphorylation on Tyr-822 is involved in the binding to

```

CC      GRB2
CC      -1- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
CC      (sialic acid binding Ig-like lectin) family.
CC      -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC      -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L16928; AAA02562.1; -
DR      EMBL; AF115401; AAD30392.1; -
DR      EMBL; AF115400; AAD30391.1; -
DR      EMBL; AF102134; AAF02417.1; -
DR      EMBL; AJ250676; CAB85609.1; -
DR      EMBL; AJ250677; CAB85610.1; -
DR      EMBL; AJ250678; CAB85611.1; -
DR      EMBL; AJ250679; CAB85612.1; -
DR      EMBL; AJ250680; CAB85613.1; -
DR      EMBL; AJ250682; CAB85615.1; -
DR      EMBL; AJ250683; CAB85616.1; -
DR      PIR; I49583; I49583.
DR      MGI; 88322; CQ22.
DR      GO; GO:0005887; C:integral to plasma membrane; IPI.
DR      GO; GO:0005515; F:protein binding; IPI.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003598; Ig_c2.
DR      Pfam; PF00047; Ig_6.
DR      SMART; SM00408; IGC2; 4.
DR      PROSITE; PS50835; IG_LIKE; 6.
KW      Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
KW      Phosphorylation; Immunoglobulin domain; Repeat; Alternative splicing.
FT      CHAIN 1 21
FT      DOMAIN 22 862
FT      TRANSMEM 703 721
FT      DOMAIN 722 862
FT      DOMAIN 147 244
FT      DOMAIN 251 341
FT      DOMAIN 346 429
FT      DOMAIN 434 515
FT      DOMAIN 520 597
FT      DOMAIN 608 691
FT      SITE 775 780
FT      SITE 835 840
FT      SITE 855 860
FT      DISULFID 41 171
FT      DISULFID 46 106
FT      DISULFID 165 229
FT      DISULFID 272 324
FT      DISULFID 368 411
FT      DISULFID 457 499
FT      DISULFID 544 586
FT      DISULFID 631 674
FT      MOD_RES 777
FT      MOD_RES 822
FT      MOD_RES 837
FT      MOD_RES 857
FT      MOD_RES 857
FT      CARBOHYD 105
FT      CARBOHYD 116
FT      CARBOHYD 139
FT      CARBOHYD 168
FT      CARBOHYD 265
FT      CARBOHYD 275

```

```

FT      CARBOHYD 378
FT      CARBOHYD 408
FT      CARBOHYD 460
FT      CARBOHYD 460
FT      CARBOHYD 561
FT      CARBOHYD 561
FT      CARBOHYD 589
FT      CARBOHYD 589
FT      VARSPIC 54
FT      VARSPIC 115
FT      VARSPIC 100
FT      VARSPIC 133
FT      VARIANT 15
FT      VARIANT 15
FT      VARIANT 19
FT      VARIANT 19
FT      VARIANT 76
FT      VARIANT 83
FT      VARIANT 90
FT      VARIANT 94
FT      VARIANT 102
FT      VARIANT 102
FT      VARIANT 112
FT      VARIANT 114
FT      VARIANT 173
FT      VARIANT 173
FT      VARIANT 179
FT      VARIANT 179
FT      VARIANT 186
FT      VARIANT 186
FT      VARIANT 190
FT      VARIANT 190
FT      VARIANT 192
FT      VARIANT 192
FT      VARIANT 236
FT      VARIANT 236
FT      VARIANT 241
FT      VARIANT 241
FT      VARIANT 244
FT      VARIANT 244

Query Match
Best Local Similarity 7.8%; Score 102.5; DB 1; Length 862;
Matches 52; Conservative 33; Mismatches 85; Indels 67; Gaps 10;

QY 32 KKGDYELTC--TASQKSIQFMWNSNQIKILGNGSFLTKGPSKLNDRADSRSLMDQ 89
DB 535 RAGQRLVLCDFASNPAPVRFWKKNGSLVDEGRYLSGVSPE-----DS 581
QY 90 GNPFLLIKKUKLEDSTYICVEVDQKEVQLV---FGLTNSPDHLLQGOSTLTLES 145
DB 582 GNYNCVNN-----SIGETLSQAMNLOVYAPRRLRVVISISPGDWVEGKATISCS 633
QY 146 ---PPGSSPVQCRSPRGKNI--QGGKTLVSQLELDGSGTWCTVQLQNGKVEFKIDIY 201
DB 634 DANPPLISQYTWPPSS--GQDLHSSGQKRLRPLEVHTSSYRC-----674
QY 202 RASALPAPPTGSALPDPOGTASALPDPPAASALPALAVISFLGLGLG---VACVIA 255
DB 675 --KGTNGIGTGESPFSTLTIVYSPS-----TIGKRVALLGFLCTICILA 717

RESULT 62
VGR2_COTUA STANDARD; PRT; 1348 AA.
ID VGR2_COTUA
AC P52583;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular endothelial growth factor receptor 2 precursor (BC 2.7.1.112)
DE (VEGFR-2) (Endothelial kinase receptor EKI) (Quek 1) (Quek1).
GN KDR OR FLK-1 OR EKI.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OC NCBI_TaxId=93934;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=97017121; PubMed=8863722;
RA Eichmann A., Marcelle C., Breant C., Le Douarin N.M.;
RT "Molecular cloning of Quek 1 and 2, two quail vascular endothelial
growth factor (VEGF) receptor-like molecules."

```

RL Gene 174:3-8(1996).

RP [2]

RC SEQUENCE OF 910-1348 FROM N.A.

RA TISSUE-Spinal cord;

RX MEDLINE=93378866; PubMed=8396413;

RT Elchmann A., Marcelle C., Breant C., Le Douarin N.M.;

RT "Two molecules related to the VEGF receptor are expressed in early

RL endothelial cells during avian embryonic development.";

RL Mech. Dev. 42:33-48(1993).

RN [3]

RP SEQUENCE OF 764-880 FROM N.A., AND CHARACTERIZATION.

RC TISSUE-Embryo;

RX MEDLINE=95301109; PubMed=7781909;

RA Flamme I., Breier G., Risau W.;

RT "Vascular endothelial growth factor (VEGF) and VEGF receptor 2 (Flk-1)

RT are expressed during vasculogenesis and vascular differentiation in

RL the adult embryo.";

RL Dev. Biol. 169:699-712(1995).

CC -1- FUNCTION: RECEPTOR FOR VEGF OR VEGF-C. HAS A TYROSINE-PROTEIN

CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM

CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF

CC VASCULAR PERMEABILITY.

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

CC tyrosine phosphate.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: IN ALL ENDOTHELIAL TISSUES DURING ONSET OF

CC VASCULARIZATION. IN LATER DEVELOPMENT, PRESENT IN LUNG, HEART,

CC INTESTINE AND SKIN.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN WHOLE MESODERM AT ONSET OF

CC GASTRULATION. FROM DAY 2, CONFINED TO ENDOTHELIAL TISSUES AND

CC EXPRESSION CONTINUES TO BE WIDESPREAD THROUGHOUT VASCULARIZATION

CC UNTIL E9 WHERE IT BECOMES RESTRICTED TO SPECIFIC REGIONS SUCH AS

CC THE SIGNAL CHORD AND HEART VALVES.

CC -1- INDUCTION: IN VITRO, VEGF IS INDUCED BY BASIC FIBROBLAST GROWTH

CC FACTOR (bFGF). UNIQUELY IN THE FIRST 24 H OF CELL CULTURE.

CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases.

CC CSF-1/PDGF receptor subfamily.

CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X83288; CAAS8268.1; -

DR EMBL; S65205; AAB28127.1; -

DR EMBL; S78345; AAB34594.1; -

DR PIR; JC4953; S51656.

DR HSP; P11362; IFGK.

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003598; IG_c2.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR001824; Recepttyrknsiii.

DR InterPro; IPR001245; Tyr_kinase.

DR InterPro; IPR008266; Tyr_kinase_AS.

DR Pfam; PF00047; Ig; 6.

DR Pfam; PF00069; Pkinase; 1.

DR ProDom; PD000001; Prot_kinase; 2.

DR SMART; SM00408; IGC2; 1.

DR SMART; SM00219; Tyrc; 1.

DR PROSITE; PS50835; IG-LIKE; 5.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.

DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.

DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.

DR Angiogenesis; Signal; Transferase; Tyrosine-protein kinase; Receptor;

KM Transmembrane; Glycoprotein; Phosphorylation; ATP-binding;

KW Immunoglobulin domain; Repeat.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 1348 VASCULAR ENDOTHELIAL GROWTH FACTOR

FT FT DOMAIN 21 756 RECEPTOR 2.

FT FT TRANSMEM 757 777 EXTRACELLULAR (POTENTIAL).

FT FT DOMAIN 778 1348 POTENTIAL.

FT FT DOMAIN 43 106 CYTOPLASMIC (POTENTIAL).

FT FT DOMAIN 138 202 IG-LIKE C2-TYPE 1.

FT FT DOMAIN 220 312 IG-LIKE C2-TYPE 2.

FT FT DOMAIN 320 405 IG-LIKE C2-TYPE 3.

FT FT DOMAIN 412 534 IG-LIKE C2-TYPE 4.

FT FT DOMAIN 540 651 IG-LIKE C2-TYPE 5.

FT FT DOMAIN 658 744 IG-LIKE C2-TYPE 6.

FT FT DOMAIN 825 1155 IG-LIKE C2-TYPE 7.

FT FT NP BIND 831 839 PROTEIN KINASE.

FT FT BINDING 859 859 ATP (BY SIMILARITY).

FT FT ACT SITE 1021 1021 ATP (BY SIMILARITY).

FT FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 666 666 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 688 688 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CARBOHYD 710 710 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT FT CONFLICT 865 865 A -> S (IN REF. 3).

SO SEQUENCE 1348 AA; 150305 MW; ASE4194A76FDSFB3 CRC64;

Query Match 7.88; Score 102.5; DB 1; Length 1348;

Best Local Similarity 22.98; Pred. No. 5.4;

Matches 49; Conservative 33; Mismatches 73; Indels 59; Gaps 12;

QY 9 HLLVLIQALLPATQGNKVVLGKGGTVLCTAGSOKSIOFHW-KNSNQIKLNGQS 67

DB 536 HVTRGELINIQPSQLTEK-----DNISLQCTADTFTEKLSWYLSLTHV--SQTP 584

QY 68 FLTKG--PSKLNDRADSRSL-----WDQGNFPIILNKLIEDSDTYICEVEDQKEE 117

DB 585 F---GGIPMPVCNKLDAQLKNTATVSVNGENTLLEILNLISLQDGGDVLCAQDKKAK 641

QY 118 VQ-LVYFGLTANSDFH-----LLQGSLLTLESPPGSSPSVOCRSFRG-----KN 162

DB 642 TORCLVHGLTVQSPPLHRLVGNLENQTTNI-----GETLEVL-C-TVNGVPPNITWFKN 694

QY 163 IQ-----GGKTLVSQLELQDSTWTC 184

DB 695 SETLFDSDGIVLKDGNKTLTIRVRKEDGGLYTC 728

RESULT 63

CEAL_MOUSE

ID CEAL_MOUSE STANDARD; PRT; 521 AA.

AC P31809;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Carcinoembryonic antigen-related cell adhesion molecule 1 precursor

DE (Biliary glycoprotein 1) (BGP-1) (Murine hepatitis virus receptor)

DE (HEV-R) (Biliary glycoprotein D).

GN CEACAM1 OR BGP OR BGP1 OR BGPD.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Colon;
 RX MEDLINE=93100765; PubMed=8380065;
 RA Dveklater G.S., Diefenbach C.B., Cardellichio C.B., McCuaig K.,
 RA Pensiero M.N.;
 RT "Several members of the mouse carcinoembryonic antigen-related
 RT glycoprotein family are functional receptors for the coronavirus
 RT mouse hepatitis virus-A59.";
 RL J. Virol. 67:1-8(1993).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RX MEDLINE=92046352; PubMed=1719235;
 RA Dveklater G.S., Pensiero M.N., Cardellichio C.B., Williams R.K.,
 RA Jiang G.-S., Holmes K.V., Diefenbach C.W.;
 RT "Cloning of the mouse hepatitis virus (MHV) receptor: expression in
 RT human and hamster cell lines confers susceptibility to MHV.";
 RL J. Virol. 65:6881-6891(1991).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93273228; PubMed=8500759;
 RA McCuaig K., Rosenberg M., Nedellec P., Turbide C., Beauchemin N.;
 RT "Expression of the Bgp gene and characterization of mouse colon
 RT biliary glycoprotein isoforms.";
 RL Gene 127:173-183(1993).
 RN (4)
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC STRAIN=CD-1; TISSUE=Colon;
 RX MEDLINE=89195121; PubMed=2702644;
 RA Beauchemin N., Turbide C., Afar D., Raymond M., Bell J.,
 RA Stanners C.P., Fuks A.;
 RT "A mouse analogue of the human carcinoembryonic antigen.";
 RL Cancer Res. 49:2017-2021(1989).
 RN (5)
 RP SEQUENCE OF 35-59.
 RX MEDLINE=91288498; PubMed=1648219;
 RA Williams R.K., Jiang G.-S., Holmes K.V.;
 RT "Receptor for mouse hepatitis virus is a member of the
 RT carcinoembryonic antigen family of glycoproteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5533-5536(1991).
 CC -1- FUNCTION: Unknown: receptor for murine coronavirus MHV-A59.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P31809-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P31809-2; Sequence=VSP_002484, VSP_002485;
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. CEA family.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
 CC or send an email to license@ebi.ac.uk).

FT SIGNAL 1 34
 FT CHAIN 35 521
 FT DOMAIN 35 428
 FT TRANSMEM 429 447
 FT DOMAIN 448 521
 FT DOMAIN 35 142
 FT DOMAIN 147 234
 FT DOMAIN 239 319
 FT DOMAIN 323 411
 FT DISULFID 167 217
 FT DISULFID 261 301
 FT DISULFID 346 394
 FT CARBOHYD 71 71
 FT CARBOHYD 89 89
 FT CARBOHYD 104 104
 FT CARBOHYD 148 148
 FT CARBOHYD 152 152
 FT CARBOHYD 199 199
 FT CARBOHYD 206 206
 FT CARBOHYD 210 210
 FT CARBOHYD 226 226
 FT CARBOHYD 258 258
 FT CARBOHYD 290 290
 FT CARBOHYD 294 294
 FT CARBOHYD 304 304
 FT CARBOHYD 317 317
 FT CARBOHYD 333 333
 FT CARBOHYD 375 375
 FT VARSPLIC 455 458
 FT VARSPLIC 459 521
 SQ SEQUENCE 521 AA; 57015 MW; 1CB871FAC47D54E CRC64;
 FTid=VSP_002485.
 FTid=VSP_002485.
 Query Match 7.7%; Score 102; DB 1; Length 521;
 Best Local Similarity 22.5%; Pred. No. 1.8;
 Matches 42; Conservative 35; Mismatches 68; Indels 42; Gaps 10;
 QY 16 LALLPATGKGVKGGDYELTCTA-SQKSIQFHKNNSQIKLQSGSFLTKGDS 74
 DB 143 ILKRPITSNNNSNPV-EGDVSILFCDSTDPDNIYLSRN-----GSLSEGD- 191
 QY 75 KLNDRADSRSLMDQGNFLIKNLKIEDSDYICEV-----DQKEVQL-LVPG----- 124
 DB 192 -----RLKISEGNRTLLNVTNRDTPGVCTRNPSVNRSDPSLNTIYGPDP 242
 QY 125 LTPNSDTHLLQGSLTLTLTLESPPGSSPSVQC-----RSPRGKNIQGGKTLVSQLELPDS 179
 DB 243 IISPSDIYIHPGSNLTLSGHA--ASNPPAQYFWLINEXPHASS---QELFIPNITNNS 296
 QY 180 GTWTCV 186
 DB 297 GTVTCV 303
 RESULT 64
 VGR2_RAT STANDARD; PRT; 1343 AA.
 AC 008775;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vascular endothelial growth factor receptor 2 precursor (EC 2.7.1.112)
 DE (VEGFR-2) (Protein-tyrosine kinase receptor flk-1) (Fetal liver kinase
 DE 1).
 GN KDR OR FLK1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.

RC TISSUE=Retina;
 RA Men Y., Edelman J.L., De Vries G.W., Sachs G.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR VEGF OR VEGF-C. HAS A TYROSINE-PROTEIN
 CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
 CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF
 CC VASCULAR PERMEABILITY (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases.
 CC CSD-1/PDGF receptor subfamily.

 CC The SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC
 DR EMBL; U93306; AAB97508.1; -;
 DR EMBL; U93307; AAB97509.1; -;
 DR HSSP; P11362; 1FCG.
 DR InterPro: IPR007110; IG-1-like.
 DR InterPro: IPR003598; PRCt_kinase.
 DR InterPro: IPR000719; PRCt_kinase.
 DR InterPro: IPR001824; ReceptTyKineIII.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR008266; Tyr_kinase_AS.
 DR Pfam; PF00047; Ig; 6.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 2.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS50835; IG_LIKE; 5.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR Angiogenesis; Transferrase; Tyrosine-protein kinase; Phosphorylation;
 KW Arg-binding; Receptor; Transmembrane; Signal; Immunoglobulin domain;
 KW Repeat; Glycoprotein.
 FT CHAIN 1 19 POTENTIAL.
 FT 20 1343 VASCULAR ENDOTHELIAL GROWTH FACTOR
 FT TRANSMEM 760 RECEPTOR 2.
 FT DOMAIN 782 POTENTIAL.
 FT DOMAIN 783 1343 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 46 109 IG-LIKE C2-TYPE 1.
 FT DOMAIN 141 207 IG-LIKE C2-TYPE 2.
 FT DOMAIN 224 320 IG-LIKE C2-TYPE 3.
 FT DOMAIN 328 414 IG-LIKE C2-TYPE 4.
 FT DOMAIN 421 540 IG-LIKE C2-TYPE 5.
 FT DOMAIN 547 654 IG-LIKE C2-TYPE 6.
 FT DOMAIN 663 749 IG-LIKE C2-TYPE 7.
 FT DOMAIN 830 1158 PROTEIN KINASE.
 FT NP_BIND 836 844 ATP (BY SIMILARITY).
 FT BINDING 864 864 ATP (BY SIMILARITY).
 FT ACT_SITE 1024 1024 BY SIMILARITY.
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 700 700 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 1055 1055 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 SQ SEQUENCE 1343 AA; 150393 MW; AD7E509B62D3FF4 CRC64;
 Query March 7.7%; Score 102; DB 1; Length 1343;
 Best Local Similarity 20.1%; Pred. No. 5.9; Indels 60; Gaps 9;
 Matches 43; Conservative 37; Mismatches 74;
 QY 9 HLLVLQALPLPAATGKNKVLGKKDPTVELCTPAGSKSIQFMKNSNQIKLGNQGSF 68
 DB 542 HVINGPEITVQPAQPTER-----ESMSLLCTADNTEFNLTWYK-----LGSQATS 588
 QY 69 LTKPSPKLNDRADSRRLM-----DQGNFPIIT--KLIKEDSTYICVEVDKBE- 117
 DB 589 VHMGES-LTPVCKNLDLMTKNGTWFNSNTDILIVAFQNASLDQGNVYCSADKKTKK 647
 QY 118 -----VQLVFGFLPANSDTLLLOGSLTLTLESPGSSPSVQCRSPRGRK----- 162
 DB 648 RHCLVKQVLIERMAPITGNLENQTTTI-----GETIEVVC--PTSGNPTPLITWFKD 699
 QY 163 -----IOGKTLVSQLELDQSGTWTC 184
 DB 700 NETLVDSGLVLDKGNRNLTIRVRKEDGGLYTC 733
 RESULT 65
 SHSI RAT STANDARD; PRT; 509 AA.
 AC P97710: 008951; 070426; GQOW15;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Protein-tyrosine phosphatase non-receptor type substrate 1 precursor
 DE (SHP substrate-1) (SHPs-1) (inhibitory receptor SHP-1) (Signal-
 DE regulatory protein alpha-1) (Sitrp-alpha-1) (Brain Ig-1-like molecule
 DE with tyrosine-based activation motifs) (Blt) (Macrophage fusion
 DE receptor) (Macrophage membrane protein MFPI50).
 GN PTPN1 OR SHPS1 OR S1RP OR BIT OR MFR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 42-60; 68-91; 128-137; 150-158;
 RP 174-189; 192-202; 204-212; 218-237; 259-270; 279-282; 405-415 AND
 RP 446-453, N-GLYCOSYLATION, PHOSPHORYLATION ON TYROSINE RESIDUES, AND
 RP INTERACTIONS WITH PTPN6 AND PTPN11.
 RC TISSUE=Fetal fibroblast;
 RX MEDLINE=97098667; PubMed=8943344;
 RA Fujioke Y., Matozaki T., Noguchi T., Yamatsu A., Yamao T.,
 RA Takahashi N., Teuda M., Takada T., Kasuga M.;
 RT "A novel membrane glycoprotein, SHP-1, that binds the SH2-domain-
 RT containing protein tyrosine phosphatase SHP-2 in response to mitogens
 RT and cell adhesion.";
 RL Mol. Cell. Biol. 16:6887-6899 (1996).
 RN [2]
 RP SEQUENCE FROM N.A., SEQUENCE OF 32-48 AND 446-453, FUNCTION, AND
 RP PHOSPHORYLATION ON TYROSINE RESIDUES.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=97415431; PubMed=9271230;
 RA Sano S.-I., Ohnishi H., Omori A., Hasegawa J., Kubota M.;
 RT "BIT, an immune antigen receptor-like molecule in the brain.";
 RL FEBS Lett. 411:327-334 (1997).
 RN [3]
 RP SEQUENCE FROM N.A., SEQUENCE OF 99-107; 128-149; 197-217; 405-417;
 RP 419-429; 446-467 AND 496-506, N-GLYCOSYLATION, AND TISSUE SPECIFICITY.
 RC STRAIN=Rhesus 344; TISSUE=Macrophage;
 RX MEDLINE=98449911; PubMed=9774638;
 RA Saginario C., Sterling H., Beckers C., Kobayashi R., Solimena M.,
 RA Ullu E., Vignery A.;
 RT "MFR, a putative receptor mediating the fusion of macrophages.";

RL M01. Cell Biol. 18:6213-6223(1998).
 RN [4]
 RN SEQUENCE OF 1-419 FROM N.A., AND TISSUE SPECIFICITY.
 RN STRAIN=MAG/RJ1; TISSUE=Alveolar macrophage;
 RC MEDLINE=98375871; PubMed=97112053;
 RA Adams S., van der Laan L.J.W., Vernon-Wilson E.,
 RA Renardel de Lavalette C., Doepp E.A., Dijkstra C.D., Simmons D.L.,
 RA van den Berg T.K.,
 RA "Signal-regulatory protein is selectively expressed by myeloid and
 RT neuronal cells.";
 RL J. Immunol. 161:1853-1859(1998).
 RN [5]
 RN PHOSPHORYLATION IN RESPONSE TO EGF, AND INTERACTION WITH PRPN1.
 RN MEDLINE=98008865; PubMed=934856;
 RP Ochi F., Matczak T., Noguchi T., Fujioke Y., Yamo T., Takeda T.,
 RA Takeda T., Matczak T., Takeda H., Fukunaga K., Noguchi T.,
 RA Tada M., Takeda H., Fukunaga K., Okabayashi Y., Kasuga M.,
 RT "Epidermal growth factor stimulates the tyrosine phosphorylation of
 RT SHP-1 and association of SHP-1 with SHP-2, a SH2 domain-containing
 RT protein tyrosine phosphatase.";
 RL Biochem. Biophys. Res. Commun. 239:483-487(1997).
 RN [6]
 RN PHOSPHORYLATION OF TYR-477 AND TYR-501, AND MUTAGENESIS OF TYR-436;
 RP TYR-460; TYR-477 AND TYR-501.
 RX MEDLINE=98204923; PubMed=9535915;
 RA Takeda T., Matczak T., Takeda H., Fukunaga K., Noguchi T.,
 RA Fujioke Y., Okazaki I., Tada M., Yamo T., Ochi F., Kasuga M.,
 RT "Roles of the complex formation of SHP-1 with SHP-2 in
 RT insulin-stimulated mitogen-activated protein kinase activation.";
 RL J. Biol. Chem. 273:9234-9242(1998).
 CC -1- FUNCTION: Immunoglobulin-like cell surface receptor for CD47. Acts
 CC as docking protein and induces translocation of PRPN6,
 CC PRPN1 and other binding partners from the cytosol to the
 CC plasma membrane. Supports adhesion of cerebellar neurons, neurite
 CC outgrowth and glial cell attachment. May play a key role in
 CC intracellular signaling during synaptogenesis and in synaptic
 CC function. Involved in the negative regulation of receptor tyrosine
 CC kinase-coupled cellular responses induced by cell adhesion, growth
 CC factors or insulin. Mediates negative regulation of phagocytosis,
 CC mast cell activation and dendritic cell activation. CD47 binding
 CC prevents maturation of immature dendritic cells and inhibits
 CC cytokine production by mature dendritic cells. May play a role in
 CC the release of nitric oxide by macrophages (By similarity).
 CC -1- SUBUNIT: Binds PRPN1 when tyrosine-phosphorylated, except in
 CC macrophages, where it primarily binds PRPN6. Binds GNB2 in vitro.
 CC Binds FGR. Binds JAK2 irrespective of its phosphorylation status
 CC and forms a stable complex. Binds SCAP1 and/or SCAP2. The
 CC resulting complex recruits FYN. Binds PRKX2 (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain, spleen, lung, liver
 CC and kidney. Detected at lower levels in heart. Highly expressed in
 CC alveolar and peritoneal macrophages, and at lower levels in
 CC dendritic cells.
 CC -1- PTM: N-glycosylated.
 CC -1- PTM: Phosphorylated on tyrosine residues in response to insulin,
 CC cell adhesion or epidermal growth factors. Dephosphorylated by
 CC PRPN1.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C1-type domains.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D85183; BAA12734.1; -;
 CC EMBL: D38468; BAA20368.1; -;
 CC EMBL: U62328; AAC68478.1; -;
 CC EMBL: AF055065; AAC18089.1; -;
 CC HSSP: P01703; 7FAB.
 CC Interpro: IPR007110; IG-like.

DR	InterPro: IPR003597; IG_c1.		
DR	InterPro: IPR003066; IG_MHC.		
DR	Pfam: Pf00047; Ig; 3.		
DR	SMART: SM00407; IGL1; 2.		
DR	PROSITE: PSS0835; IG_Like; 3.		
DR	PROSITE: PS00290; IG_MHC; 1.		
KW	Repeat; Signal; Transmembrane;		
KW	Glycoprotein; Phosphorylation.		
FT	SIGNAL	1	31
FT	CHAIN	32	509
FT			
FT	DOMAIN	28	373
FT	TRANSMEM	374	394
FT	DOMAIN	395	509
FT	DOMAIN	32	138
FT	DOMAIN	150	248
FT	DOMAIN	255	349
FT	DISULFID	55	122
FT	DISULFID	172	229
FT	DISULFID	274	332
FT	SITE	436	439
FT	SITE	446	451
FT	SITE	460	463
FT	SITE	477	480
FT	SITE	504	504
FT	MOD_RES	436	436
FT			
FT	MOD_RES	460	460
FT			
FT	MOD_RES	477	477
FT	MOD_RES	501	501
FT	CARBOHYD	54	54
FT	CARBOHYD	93	93
FT	CARBOHYD	169	169
FT	CARBOHYD	181	181
FT	CARBOHYD	205	205
FT	CARBOHYD	209	209
FT	CARBOHYD	242	242
FT	CARBOHYD	246	246
FT	CARBOHYD	271	271
FT	CARBOHYD	293	293
FT	CARBOHYD	312	312
FT	CARBOHYD	320	320
FT	CARBOHYD	345	345
FT	CARBOHYD	436	436
FT			
FT	MUTAGEN	460	460
FT			
FT	MUTAGEN	477	477
FT			
FT	MUTAGEN	501	501
FT			
FT	CONFLICT	8	8
FT	CONFLICT	10	10
FT	CONFLICT	25	25
FT	CONFLICT	58	58
FT	CONFLICT	99	100
FT	CONFLICT	162	162
FT	CONFLICT	189	189
FT	CONFLICT	205	205
FT	CONFLICT	209	209
FT	CONFLICT	405	405
FT	CONFLICT	416	416
FT	CONFLICT	418	421
FT	CONFLICT	450	450

[illegible]

```

CC -I- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=3;
CC Name=N-CAM 120;
CC IsoId=P13594-1; Sequence=Displayed;
CC Name=N-CAM 180;
CC IsoId=P13595-1; Sequence=External;
CC Name=N-CAM 140;
CC IsoId=P13595-2; Sequence=External;
CC -I- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -I- SIMILARITY: Contains 2 fibronectin type III domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC -----
DR EMBL; Y00051; CAA68263.1; -.
DR EMBL; X15049; CAA33148.1; ALT_SEQ.
DR EMBL; X07195; CAA30173.1; -.
DR PIR; A29673; IJMSNG.
DR PDB; 2NCM; 12-MAR-97.
DR PDB; 3NCM; 23-JUL-99.
DR MGD; MGI:97281; Ncaml.
DR InterPro; IPRO08957; FN_III-like.
DR InterPro; IPRO03961; FN_III.
DR InterPro; IPRO07110; Ig-like.
DR InterPro; IPRO03598; Ig_c2.
DR Pfam; PF00041; fn3; 2.
DR SMART; SMO0047; Ig; 5.
DR SMART; SMO0060; FN3; 2.
DR PROSITE; PS50835; IG_LIKE; 5.
DR Cell adhesion; Glycoprotein; Repeat; Alternative splicing;
KW Immunoglobulin domain; Signal; Heparin-binding; GPI-anchor;
KW 3D-structure.
FT FT SIGNAL 1 19
FT CHAIN 20 725
FT FT NEURAL CELL ADHESION MOLECULE 1, 120 kDa
FT FT ISOFORM, C2-TYPE 1.
FT DOMAIN 20 111 IG-LIKE C2-TYPE 1.
FT DOMAIN 116 205 IG-LIKE C2-TYPE 2.
FT DOMAIN 212 302 IG-LIKE C2-TYPE 3.
FT DOMAIN 309 402 IG-LIKE C2-TYPE 4.
FT DOMAIN 407 492 IG-LIKE C2-TYPE 5.
FT DOMAIN 519 596 FIBRONECTIN TYPE-III 1.
FT DOMAIN 625 692 FIBRONECTIN TYPE-III 2.
FT DOMAIN 152 156 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 161 165 HEPARIN-BINDING (POTENTIAL).
FT DISULFID 41 96 PROBABLE.
FT DISULFID 139 189 PROBABLE.
FT DISULFID 235 288 PROBABLE.
FT DISULFID 330 386 PROBABLE.
FT DISULFID 427 480 PROBABLE.
FT CARBOHYD 316 316 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 424 424 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 479 479 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 261 268 VERSSAYS -> DEKHSFD (IN REF. 2).
FT CONFLICT 273 273 V -> L (IN REF. 2).
FT CONFLICT 354 355 OD -> KT (IN REF. 2).
FT CONFLICT 549 549 T -> K (IN REF. 2).
FT CONFLICT 572 572 T -> R (IN REF. 2).
FT CONFLICT 575 575 D -> V (IN REF. 2).
FT CONFLICT 589 594 MOSES -> SATER (IN REF. 2).
FT CONFLICT 600 602 PEL -> REP (IN REF. 2).
FT CONFLICT 657 657 H -> D (IN REF. 2 AND 3).
SQ SEQUENCE 725 AA; 80296 MW; C2AE8B846IC6BZF CRC64;

```

Best Local Similarity 19.9%; Pred. No. 3;
Matches 46; Conservative 42; Mismatches 94; Indels 49; Gaps 10;

QY 3 RG-VPRHLLVLQALLPAATGKRVLGKGGDTVELTCTASOKSIOFHW-KNSNOIK 60
196 RGEINFDIOIVNVPPVQARQSVNATANGQSVTLVCDADGPFPEPTMSWTKDSEPI 255

QY 61 ILGNQGSFLTKGSKLNDRAD--SRSLMIOGNFPLIKLKIEDSDTYICEVE---DQ 114
256 -----NEEDERSRVSQSSSE--VIRRVNDKDEAEVCIENKAGEQ 297

QY 115 KEEVQLLVFG---LTANSDTHLLQGSLLTLTLESPGSSPSVQCR-----SPRGNIQ 164
298 DASHLNVFAKPKITYVENQTMLEEQVTLTCEASDPIPSITWRSTNISEEQDL 357

QY 165 GG-----KTLVSQLELDSDGTWC---TVLQNKVKFEKIDIVER 202
358 GHMVRSHARVSSLTKSIQYRDAGEVWCTASNTIGDSQSIDLEFGYAPK 408

Db 358 GHMVRSHARVSSLTKSIQYRDAGEVWCTASNTIGDSQSIDLEFGYAPK 408

RESULT 67
NCAL_MOUSE STANDARD; PRT; 858 AA.

AC P13595;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 140 kDa isoform precursor (N-CAM 140)
DE (NCAM-140).
GN NCAM1 OR NCAM.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=80059265; PubMed=1680385;
RA Small S.O., Shull G.E., Santoni M.-J., Akesson R.;
RT "Identification of a cDNA clone that contains the complete coding
sequence for a 140-kD rat NCAM polypeptide.";
RL J. Cell Biol. 105:2335-2345(1987).
RN [2]
RP SEQUENCE OF 355-364 FROM N.A.
RX MEDLINE=90166485; PubMed=2483093;
RA Small S.O., Haines S.L., Akesson R.A.;
RT "Polypeptide variation in an N-CAM extracellular immunoglobulin-like
RT fold is developmentally regulated through alternative splicing.";
RL Neuron 11:1007-1017(1988).
CC -1- FUNCTION: This protein is a cell adhesion molecule involved in
CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced;
CC Name=1;
CC IsoId=P13596-1; Sequence=displayed;
CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06564; CAA29609.1; -;
DR EMBL; M32611; AAA1679.1; -;
DR PIR; S00846; IJRTNC.
DR PDB; 1EPF; 27-OCT-00.

DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; IG_1-like.
DR InterPro; IPR003598; IG_C2.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGC2; 5.
DR PROSITE; PS00835; IG_LIKE; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Repeat;
KW Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding;
KW 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 858
FT DOMAIN 20 721
FT TRANSMEM 722 739
FT DOMAIN 740 858
FT DOMAIN 20 111
FT DOMAIN 116 205
FT DOMAIN 212 302
FT DOMAIN 309 414
FT DOMAIN 417 502
FT DOMAIN 514 615
FT DOMAIN 616 712
FT DOMAIN 152 165
FT DOMAIN 161 185
FT DISULFID 41 96
FT DISULFID 139 189
FT DISULFID 235 288
FT DISULFID 330 396
FT DISULFID 437 490
FT CARBOHYD 222 222
FT CARBOHYD 316 316
FT CARBOHYD 348 348
FT CARBOHYD 434 434
FT CARBOHYD 460 460
FT CARBOHYD 489 489
SQ SEQUENCE 858 AA; 94658 MW; EALN064AE050F6 CRC64;
Query Match 7.7%; Score 101.5; DB 1; Length 858;
Best Local Similarity 18.0%; Pred. No. 3.7;
Matches 43; Conservative 42; Mismatches 99; Indels 55; Gaps 8;

QY 3 RG-VPRHLLVLQALLPAATGKRVLGKGGDTVELTCTASOKSIOFHW-KNSNOIK 60
196 RGEINFDIOIVNVPPVQARQSVNATANGQSVTLVCDADGPFPEPTMSWTKDSEPI 255

QY 61 ILGNQGSFLTKGSKLNDRAD--SRSLMIOGNFPLIKLKIEDSDTYICEVE---DQ 116
256 -----NEEDERSRVSQSSSE--VIRRVNDKDEAEVCIENKAGEQ 299

QY 117 EVQLLVFG---LTANSDTHLLQGSLLTLTLESPGSSPSVQCR-----DQ 157
300 SHLKVFAPKPKITYVENQTMLEEQVTLTCEASDPIPSITWRSTNISEKASWTR 359

QY 158 PRKANIQCG-----KTLVSQLELDSDGTWC---TVLQNKVKFEKIDIVER 202
360 PEKQETLDGHMVRSHARVSSLTKSIQYRDAGEVWCTASNTIGDSQSVTLVQYAPK 418

Db 360 PEKQETLDGHMVRSHARVSSLTKSIQYRDAGEVWCTASNTIGDSQSVTLVQYAPK 418

RESULT 68
NCAL_MOUSE STANDARD; PRT; 1115 AA.

AC P13595; Q61949;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM 180)
DE (NCAM-180).
GN NCAM1 OR NCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM N-CAM 180).
 RC STRAIN=C57BL/6;
 RX MEDLINE=87246524; Pubmed=3595563;
 RA Barthels D., Santoni M.J., Wille W., Ruppert C., Chaix J.C.,
 Hirsch M.R., Fontecilla-Camps J.C., Goridis C.;
 RT "Isolation and nucleotide sequence of mouse NCM cDNA that codes for
 a Mr 79,000 polypeptide without a membrane-spanning region.";
 RL EMBO J. 6:907-914(1987).
 RN [2]
 RP SEQUENCE OF 529-1115 FROM N.A. (ISOFORM N-CAM 140).
 RC STRAIN=C57BL/6;
 RX MEDLINE=88067687; Pubmed=3684567;
 RA Santoni M.-J., Barthels D., Barbas J.A., Hirsch M.-R., Steinmetz M.,
 Goridis C., Wille W.;
 RT "Analysis of cDNA clones that code for the transmembrane forms of the
 mouse neural cell adhesion molecule (NCAM) and are generated by
 alternative RNA splicing.";
 RL Nucleic Acids Res. 15:8621-8641(1987).
 RN [3]
 RP SEQUENCE OF 642-1115 FROM N.A. (ISOFORM N-CAM 180).
 RX MEDLINE=88283628; Pubmed=3396534;
 RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;
 RT "Differential splicing and alternative polyadenylation generates
 distinct NCAM transcripts and proteins in the mouse.";
 RL EMBO J. 7:625-632(1988).
 RN [4]
 RP SEQUENCE OF 804-1081 FROM N.A. (ISOFORM N-CAM 180).
 RC STRAIN=C57BL/6J; Tissue=Brain;
 RX MEDLINE=88247737; Pubmed=2454455;
 RA Barthels D., Vopper G., Wille W.;
 RT "NCAM-180, the large isoform of the neural cell adhesion molecule of
 the mouse, is encoded by an alternatively spliced transcript.";
 RL Nucleic Acids Res. 16:4217-4225(1988).
 RN [5]
 RP SEQUENCE OF 702-1115 FROM N.A. (ISOFORM N-CAM 140).
 RC STRAIN=C57BL/6; Tissue=Brain;
 RX MEDLINE=89251563; Pubmed=2721486;
 RA Santoni M.J., Barthels D., Vopper G., Boned A., Goridis C., Wille W.;
 RT "Differential exon usage involving an unusual splicing mechanism
 generates at least eight types of NCAM cDNA in mouse brain.";
 RL EMBO J. 8:385-392(1989).
 RN [6]
 RP SEQUENCE OF 20-36.
 RX MEDLINE=86140120; Pubmed=3512556;
 RA Rougon G., Marshak D.R.;
 RT "Structural and immunological characterization of the amino-terminal
 domain of mammalian neural cell adhesion molecules.";
 RL J. Biol. Chem. 261:3396-3401(1986).
 RL J. Biol. Chem. 261:3396-3401(1986).
 CC -1- FUNCTION: This protein is a cell adhesion molecule involved in
 neuron-neuron adhesion, neurite fasciculation, outgrowth of
 neurites, etc.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=N-CAM 180;
 CC IsoId=P13595-1; Sequence=Displayed;
 CC Name=N-CAM 140;
 CC IsoId=P13595-2; Sequence=VSP_002588;
 CC Name=N-CAM 120;
 CC IsoId=P13594-1; Sequence=External;
 CC -1- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 or send an email to license@ebi.ac.uk).

CC -----
 DR EMBL; X07200; CAA30177.1; -;
 DR EMBL; X00051; -; NOT ANNOTATED_CDS.
 DR EMBL; X06328; CAA29641.1; -;
 DR EMBL; X07195; CAA30173.1; -;
 DR EMBL; X07244; CAA30230.1; -;
 DR EMBL; X15051; CAA33150.1; -;
 DR EMBL; X15052; CAA33151.1; -;
 DR PIR; A29673; IJMSNL.
 DR MGI; 97281; Ncam1.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR007110; Ig_III.
 DR InterPro; IPR003598; Ig_c2.
 DR Pfam; PF00047; fn3; 2.
 DR Pfam; PF00047; Ig; 5.
 DR SMART; SM00060; FN3; 2.
 DR SMART; SM00408; IGC2; 5.
 DR PROSITE; PS50835; IG_LIKE; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Repeat;
 KW Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding.
 FT SIGNAL 1 19
 FT CHAIN 20 1115
 FT FT
 FT DOMAIN 20 711 NEURAL CELL ADHESION MOLECULE 1, 180 kDa
 FT TRANSMEM 712 729 ISOFORM.
 FT DOMAIN 730 1115 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 20 111 POTENTIAL.
 FT DOMAIN 116 205 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 212 302 IG-LIKE C2-TYPE 1.
 FT DOMAIN 309 402 IG-LIKE C2-TYPE 2.
 FT DOMAIN 407 492 IG-LIKE C2-TYPE 3.
 FT DOMAIN 519 596 IG-LIKE C2-TYPE 4.
 FT DOMAIN 625 692 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 152 156 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 161 165 HEPARIN-BINDING (POTENTIAL).
 FT DISULFID 41 96 HEPARIN-BINDING (POTENTIAL).
 FT DISULFID 139 188 PROBABLE.
 FT DISULFID 235 288 PROBABLE.
 FT DISULFID 330 386 PROBABLE.
 FT CARBOHYD 427 480 PROBABLE.
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 810 1076 Missing (in isoform N-CAM 140).
 FT /FTid=VSP_002588.
 SQ SEQUENCE 1115 AA; 119351 MW; 2C93DCD474CFBCAF CRC64;
 Query Match 7.74; Score 101.5; DB 1; Length 1115;
 Best Local Similarity 19.94; Pred. No. 5.1;
 Matches 46; Conservative 42; Mismatches 94; Indels 49; Gaps 10;
 QY 3 RG-VPRPHLLVQLALPAPATQGNKVVLGKGDYVLTCTAAGKSIQTHW-KNSNQIK 60
 DB 196 RGEINFDIIVIVVNPPTVQAROSIVNATNLQOSVTLVCDAGFPPTMSWTKDGPTE 255
 QY 61 ILNGQSFLLTKGSPKNDRAD--SRRLMDQGNPPLIKULKIEDSTYCEV-----DQ 114
 DB 256 -----NEEDERSRSVSQSSS--VTRVNDKNDAEVCIENKAGBQ 297
 QY 115 KEEOVLIVFG---LTRANSOTHLLOQSLTLTLESPPGSSPVQCR-----SPRGKNIQ 164
 DB 298 DASHLKVFAPKPKITYVENQTAMELEQVTLTCEASGDPISITWRSTRNISSEQDLD 357
 QY 165 GG-----KTLVSQLELDQSGTWTC---TVLQNKQKVFPKIDIVR 202
 DB 358 GHWVSRHARVSSITLTKSIQYRDAGEYMCYASNTIGQDSQSIDLEFOYAPK 408
 RESULT 69

[illegible]

CC	NCBI_TaxId=9031;
RN	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC	TISSUE=Brain;
RX	MEDLINE=92174898; PubMed=1311675;
RA	Zuellig R.A., Rader C., Schroeder A., Kalousek M.B.,
RV	von Bohlen Und Halbach F., Osterwalder T., Inan G., Stoeckli E.T.,
RA	Affeller H.-U., Fritz A., Hafen E., Sonderegger P.;
RT	"Type axonally secreted cell adhesion molecule, axonin-1. Primary
RT	structure, immunoglobulin-like and fibronectin-type-III-like domains
RT	and glycosyl-phosphatidylinositol anchorage.";
RL	Eur. J. Biochem. 204:453-463(1992).
CC	-1 FUNCTION: Axon-associated cell adhesion molecule (AxCAM) which
CC	promotes neurite outgrowth by interaction with the AxCAM II (G4)
CC	of neuritic membrane.
CC	-1 SUBCELLULAR LOCATION: Attached to the neuronal membrane by a
CC	GPI-anchor.
CC	-1 PTM: The N-terminus is blocked.
CC	-1 SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC	-1 SIMILARITY: Contains 4 fibronectin type III domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: X63101; CAA44815.1; .
DR	PIR: S22383; S22383.
DR	PDB: 1CS6; 1S-MAY-00.
DR	InterPro: IPR008957; FN_III-like.
DR	InterPro: IPR003961; FN_III.
DR	InterPro: IPR007110; IG-like.
DR	InterPro: IPR003598; IG_c2.
DR	Pfam: PF00041; fn3; 3.
DR	Pfam: PF00047; Ig; 6.
DR	SMART; SM00600; FN3; 4.
DR	SMART; SM00408; IG3; 5.
DR	PROSITE: PS50835; IG_LIKE; 6.
KW	Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
KW	Cell adhesion; Repeat; 3D-structure.
FT	SIGNAL 1 23 OR 25 (POTENTIAL).
FT	CHAIN ? 1036 CONTACTIN 2.
FT	PROPEP ? 1036 REMOVED IN MATURE FORM.
FT	DOMAIN 32 123 IG-LIKE C2-TYPE 1.
FT	DOMAIN 128 223 IG-LIKE C2-TYPE 2.
FT	DOMAIN 234 317 IG-LIKE C2-TYPE 3.
FT	DOMAIN 322 406 IG-LIKE C2-TYPE 4.
FT	DOMAIN 412 499 IG-LIKE C2-TYPE 5.
FT	DOMAIN 504 598 IG-LIKE C2-TYPE 6.
FT	DOMAIN 599 608 HINGE (POTENTIAL).
FT	DOMAIN 601 607 GLY/PRO-RICH.
FT	DOMAIN 608 709 FIBRONECTIN TYPE-III 1.
FT	DOMAIN 710 811 FIBRONECTIN TYPE-III 2.
FT	DOMAIN 812 912 FIBRONECTIN TYPE-III 3.
FT	DOMAIN 913 1009 FIBRONECTIN TYPE-III 4.
FT	CARBOHD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 770 770 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 900 900 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHD 914 914 N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE 1036 AA; 113301 MW; 08B80143BE779794 CRC64;
Query Match	7.7%; Score 101; DB 1; Length 1036;
Best Local Similarity	23.2%; Pred. No. 5.1;
Matches 53; Conservative 30; Mismatches 87; Indels 58; Gaps 9;	

```

Oy 34 GDTVELTCTASQKSIQFMKNSNQIKLNGSFLTKGPKSLNDRADRSRLMDQGNFP 93
Db 249 GOMWTLCEFPAGNPPVQIKMR-----KLDGQ-----TSKMLSSPL 285
Oy 94 LIKMLKIEDSDTYICEVEDQK---EEVQLIVFG---LRNSPTHLQOSLTITLES 145
Db 286 LHIQNVDEDEBGTVECEAENIKGRDYGRIIHHQPMVLVITDTEADISDLRMSCVA 345
Oy 146 PFGSSPSYQ-----CRSPRGKNTGGKTLVSQLELSDSGTMTCTYLQNKQKV----- 193
Db 346 SGKRPAPARWLKDGOPLASQNRIVSGS-LRFSCLVEDSGMTCQVAENKRGTVYASAE 404
Oy 194 -----EFKIDIVPRASALPAPPTGSALPDPTQSALPDPPASAL 233
Db 405 LTVQALAPDFRLNPVKR--LIPARNSGKVIIPCQBRAR-----PATVTL 446

RESULT 71
UN52_CABEEL STANDARD: PRT: 3375 AA.
AC 006561: 018261; 018263; 09XTD2; 09XT15;
DT 01-JUN-1994 (Rel. 29, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane proteoglycan precursor (Perlecan homolog)
DE (Uncoordinated protein 52).
DE UNC-52 OR ZC101.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A. (ISOPROM A), AND FUNCTION.
RX MEDLINE=9339574; PubMed=893416;
RA RogalSKI T.M., Williams B.D., Mullen G.P., Moerman D.G.;
RT "Products of the unc-52 gene in Caenorhabditis elegans are homologous
RT to the core protein of the mammalian basement membrane heparan
RT sulfate proteoglycan."
RL Genes Dev. 7:1471-1484(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Percy C.M., Baynes C.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Durbin R.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Probable role in myofibril assembly and/or attachment
CC of the myofibril lattice to the cell membrane. May be an
CC extracellular anchor for integrin receptors in muscle.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS:
CC Name=;
CC Event=Alternative splicing; Named isoforms=4;
CC Name=;
CC IsoId=Q06561-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=;
CC IsoId=Q06561-2; Sequence=VSP_007195, VSP_007196;
CC Name=;
CC IsoId=Q06561-3; Sequence=VSP_007191, VSP_007192;
CC Note=No experimental confirmation available;
CC Name=;
CC IsoId=Q06561-4; Sequence=VSP_007193, VSP_007194, VSP_007195,
CC VSP_007196;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Found in the basement membrane of all
CC contractile tissues. It is concentrated over muscle dense bodies
CC and M-lines which are associated with beta-integrin.
CC -1- DEVELOPMENTAL STAGE: Synthesized early in embryogenesis.
CC -1- SIMILARITY: Contains 3 LDL-receptor class A domains.
CC -1- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 7 laminin EGF-like domains.

```

```

CC -1- SIMILARITY: Contains 3 laminin G-like domains.
CC -1- SIMILARITY: Contains 2 laminin IV domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ch/announce/
CC or send an email to license@ebi.ebi.ch).
CC -----
DR EMBL; L13458; AAA28156.1; -
DR EMBL; Z93375; CAB07567.1; -
DR EMBL; Z93395; CAB07567.1; JOINED.
DR EMBL; Z93375; CAB07568.1; -
DR EMBL; Z93395; CAB07568.1; JOINED.
DR EMBL; Z93375; CAB07569.1; -
DR EMBL; Z93395; CAB07569.1; JOINED.
DR EMBL; Z93395; CAB07704.1; -
DR EMBL; Z93395; CAB07704.1; JOINED.
DR EMBL; Z93395; CAB07706.1; -
DR EMBL; Z93375; CAB07706.1; JOINED.
DR EMBL; Z93395; CAB07707.1; -
DR EMBL; Z93375; CAB07707.1; JOINED.
DR EMBL; Z93375; CAB07707.1; JOINED.
DR EMBL; Z93375; CAB07708.1; -
DR EMBL; Z93375; CAB07708.1; JOINED.
DR HSSP; P01130; ILDR.
DR WormPep; ZC101.2a; CE15028.
DR WormPep; ZC101.2b; CE15030.
DR WormPep; ZC101.2c; CE15034.
DR WormPep; ZC101.2e; CE18424.
DR GO; GO:0005578; C:extracellular matrix; IEP.
DR GO; GO:0030239; P:myofibril assembly; IEP.
DR InterPro; IPR008985; CoA_Like_Iec_g1.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; Ig_Like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR Pfam; PF00047; Ig_16.
DR Pfam; PF00052; Laminin_B_2.
DR Pfam; PF00053; Laminin_EGF_5.
DR Pfam; PF00057; Idl_recept_a; 3.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRODOM; PD003031; Laminin_B_2.
DR SMART; SM00181; EGF_6.
DR SMART; SM00179; EGF_CA_2.
DR SMART; SM00180; EGF_Lam; 6.
DR SMART; SM00409; IG_17.
DR SMART; SM00408; IGC2_17.
DR SMART; SM00281; Lamb; 1.
DR SMART; SM00282; Lamb; 3.
DR SMART; SM00192; LDLa; 3.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00026; EGF_3; 3.
DR PROSITE; PS00835; IG_Like; 17.
DR PROSITE; PS00025; LAM_G_DOMAIN; 3.
DR PROSITE; PS01246; LAMININ_TYPE_EGF_7.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
KW Proteoglycan; Glycoprotein; Signal; Immunoglobulin domain; Repeat;
KW Basement membrane; Extracellular matrix; Alternative splicing;
KW Laminin EGF-like domain.
FT SIGNAL 1 22
FT CHAIN 23 3375
FT DOMAIN 45 130
FT DOMAIN 148 184
FT DOMAIN 189 225

```


FT	DOMAIN	232	269	LDL-RECEPTOR CLASS A 3.
FT	DOMAIN	271	355	IG-LIKE C2-TYPE 2.
FT	DOMAIN	384	431	LAMININ EGF-LIKE 1. (INCOMPLETE).
FT	DOMAIN	432	441	LAMININ EGF-LIKE 2. (N-TERMINAL).
FT	DOMAIN	442	633	LAMININ DOMAIN IV 1.
FT	DOMAIN	634	666	LAMININ EGF-LIKE 2. (C-TERMINAL).
FT	DOMAIN	674	720	LAMININ EGF-LIKE 3. (INCOMPLETE).
FT	DOMAIN	721	730	LAMININ EGF-LIKE 4. (N-TERMINAL).
FT	DOMAIN	731	921	LAMININ DOMAIN IV 2.
FT	DOMAIN	922	954	LAMININ EGF-LIKE 4. (C-TERMINAL).
FT	DOMAIN	955	1004	LAMININ EGF-LIKE 5.
FT	DOMAIN	1011	1060	LAMININ EGF-LIKE 6.
FT	DOMAIN	1061	1111	LAMININ EGF-LIKE 7.
FT	DOMAIN	1126	1222	IG-LIKE C2-TYPE 3.
FT	DOMAIN	1226	1311	IG-LIKE C2-TYPE 4.
FT	DOMAIN	1319	1401	IG-LIKE C2-TYPE 5.
FT	DOMAIN	1410	1499	IG-LIKE C2-TYPE 6.
FT	DOMAIN	1503	1585	IG-LIKE C2-TYPE 7.
FT	DOMAIN	1588	1680	IG-LIKE C2-TYPE 8.
FT	DOMAIN	1690	1785	IG-LIKE C2-TYPE 9.
FT	DOMAIN	1793	1878	IG-LIKE C2-TYPE 10.
FT	DOMAIN	1886	1970	IG-LIKE C2-TYPE 11.
FT	DOMAIN	1973	2069	IG-LIKE C2-TYPE 12.
FT	DOMAIN	2073	2163	IG-LIKE C2-TYPE 13.
FT	DOMAIN	2173	2260	IG-LIKE C2-TYPE 14.
FT	DOMAIN	2263	2343	IG-LIKE C2-TYPE 15.
FT	DOMAIN	2349	2435	IG-LIKE C2-TYPE 16.
FT	DOMAIN	2446	2530	IG-LIKE C2-TYPE 17.
FT	DOMAIN	2532	2713	LAMININ G-LIKE 1.
FT	DOMAIN	2793	2960	LAMININ G-LIKE 2.
FT	DOMAIN	2961	3093	GLU-RICH.
FT	DOMAIN	2972	3066	THR-RICH.
FT	DOMAIN	3180	3359	LAMININ G-LIKE 3.
FT	DISULFID	66	114	BY SIMILARITY.
FT	DISULFID	149	161	BY SIMILARITY.
FT	DISULFID	156	174	BY SIMILARITY.
FT	DISULFID	168	183	BY SIMILARITY.
FT	DISULFID	190	202	BY SIMILARITY.
FT	DISULFID	197	215	BY SIMILARITY.
FT	DISULFID	209	224	BY SIMILARITY.
FT	DISULFID	233	246	BY SIMILARITY.
FT	DISULFID	240	259	BY SIMILARITY.
FT	DISULFID	253	268	BY SIMILARITY.
FT	DISULFID	955	964	BY SIMILARITY.
FT	DISULFID	957	971	BY SIMILARITY.
FT	DISULFID	974	983	BY SIMILARITY.
FT	DISULFID	986	1002	BY SIMILARITY.
FT	DISULFID	1011	1021	BY SIMILARITY.
FT	DISULFID	1013	1027	BY SIMILARITY.
FT	DISULFID	1030	1039	BY SIMILARITY.
FT	DISULFID	1042	1058	BY SIMILARITY.
FT	DISULFID	1061	1069	BY SIMILARITY.
FT	DISULFID	1063	1079	BY SIMILARITY.
FT	DISULFID	1082	1091	BY SIMILARITY.
FT	DISULFID	1094	1109	BY SIMILARITY.
FT	DISULFID	1152	1200	BY SIMILARITY.
FT	DISULFID	1338	1384	BY SIMILARITY.
FT	DISULFID	1435	1481	BY SIMILARITY.
FT	DISULFID	1527	1573	BY SIMILARITY.
FT	DISULFID	1618	1663	BY SIMILARITY.
FT	DISULFID	1719	1767	BY SIMILARITY.
FT	DISULFID	1814	1861	BY SIMILARITY.
FT	DISULFID	1907	1954	BY SIMILARITY.
FT	DISULFID	1998	2053	BY SIMILARITY.
FT	DISULFID	2099	2147	BY SIMILARITY.
FT	DISULFID	2195	2242	BY SIMILARITY.
FT	DISULFID	2284	2329	BY SIMILARITY.
FT	DISULFID	2374	2420	BY SIMILARITY.
FT	DISULFID	2467	2514	BY SIMILARITY.
FT	CARBOHYD	1422	1452	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	2476	2476	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	2950	2950	N-LINKED (GLCNAC. . .)
FT	CARBOHYD	3143	3143	N-LINKED (GLCNAC. . .)

Query Match	Best Local Similarity	7.7%;	Score 101;	DB 1;	Length 3375;
Matches 50;	Conservative 37;	Mismatches 106;	Indels 76;	Gaps 10;	
Qy	2	NRGVPFRLHLLVQLALPLPAATGKGNKVVLGKKGDVLELTCTASQSKSIQFMKNNSQIKI	61		
Db	1581	NRPEVSPNPAIVYKSPRPRIIDPAEQTV--PESSPKIRKYVGHBSVQLTFRRVS----	1634		
Qy	62	LGNQGSFLTKGPKSLNDRADSRSLWDQGNFPIIKNLKIEDSDTYICEVED-----Q	114		
Db	1635	-----GQLNEDADENNGL-----LAVQRAELTDGSDYICTARDPDTCAPID	1675		
Qy	115	KEEVQVLLVPELTANSPTDHLQGSLLTILES---PEGSSSVQCRSP-----	158		
Db	1676	STPATVYNTAAAPPOVEARPPHPVITTPQTTIPGSDPARICTVYGNPSAAQHSFER	1735		
Qy	159	-RGNNIQGKT-----LSVQLLEDSGTCTVVLGNQKVEFKIDIVPRASALPAPTG	212		
Db	1736	VDKGLEPFSSDDRGVLTISTQLQDAGEVCLY-----SPENSPPVK	1778		
Qy	213	SALPDPTQASALPD--PPASALPALAV	239		
Db	1779	T--NPSTLNTPEGTPPVPATPPLISV	1804		
RESULT 72					
KV5Q_MOUSE					
ID	KV5Q_MOUSE	STANDARD;	PRT;	108 AA.	
AC	P01650;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	21-JUL-1986 (Rel. 01, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	IG kappa chain V-V region UPC 61.				
DS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euleleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE.				
RX	MEDLINE=79195288; PubMed=109517;				
RA	Viana M., Rudikoff S., Potter M.;				
RT	"The structural basis of a hapten-inhibitable kappa-chain idotype";				
RL	J. Immunol. 122:1905-1910(1979)				
CC	-I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT				
CC	BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INDULIN).				
CC	PIR; A92808; KVM561.				
DR	HSSP; P80362; 1MTL.				
DR	InterPro; IPR007110; IG_1like.				
DR	InterPro; IPR003596; IG_V.				
DR	Pfam; PF00047; IG; 1.				
DR	SMART; SM00406; IGv; 1.				
DR	PROSITE; PSS0835; IG_LIKE; 1.				
KW	Immunoglobulin V region.				
FT	DOMAIN	1	23	FRAMEWORK-1.	
FT	DOMAIN	24	34	COMPLEMENTARITY-DETERMINING-1.	
FT	DOMAIN	35	49	FRAMEWORK-2.	
FT	DOMAIN	50	56	COMPLEMENTARITY-DETERMINING-2.	
FT	DOMAIN	57	88	FRAMEWORK-3.	
FT	DOMAIN	89	97	COMPLEMENTARITY-DETERMINING-3.	
FT	DOMAIN	98	108	FRAMEWORK-4.	
FT	DISULFID	23	88	BY SIMILARITY.	
FT	NON TER	108	108		
SEQ	SEQUENCE	108 AA;	11809 MW;	FAEADA36076F2AFE CRC64;	
Query Match	Best Local Similarity	7.6%;	Score 100.5;	DB 1;	Length 108;
Matches 30;	Conservative 5;	Mismatches 37;	Indels 5;	Gaps 2;	
Qy	34	GDTVELTCTASQKSKSIQFMKNNSQIKILGNQGSFLTKGPKSLNDRADSRSLWDQ--NF	92		
Db	16	GDIYTMTCQASQGSITIMNFWQKRP--GKAPLTLIGASILEDGVPFRFSGSRGTDGF	71		

OY 93 PLINKLKIEDSTYIC 109
 DB 72 TLTIISLEDEDMATYFC 88

RESULT 73
 OPCM_CHICK STANDARD; PRT; 337 AA.

AC 098852;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Opioid binding protein/cell adhesion molecule homolog precursor
 GN (Neurite inhibitor GP55-A) (OBAM protein gamma isoform).
 OS Opioid.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Lodge A.P., Kim D.-S., Howard M.R., McNamee C.J., Smith N., Moss D.J.;
 RT "Cloning of CEPU-8, a secreted isoform of CEPU-1, and OBAM CDNA from
 RT chick: structural diversity of IGLON family proteins.";
 RL Submitted (Apr-1999) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 64-337 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=97157768; PubMed=9004047;
 RA Wilson D.J.A., Kim D.-S., Clarke G.A., Marshall-Clarke S., Moss D.J.;
 RT "A family of glycoproteins (GP55), which inhibit neurite outgrowth,
 RT are members of the Ig superfamily and are related to OBAM,
 RT neurotrophin, LAMP and CEPU-1.";
 RL J. Cell Sci. 109:3129-3138(1996).
 CC -1- FUNCTION: Inhibits neurite outgrowth.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: Restricted to the nervous system.
 CC -1- DEVELOPMENTAL STAGE: Increases during development from very low
 CC levels at embryonic day 10 and is most abundant after hatching.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. IGLON
 CC family.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y08170; CAB41420.1; -
 CC InterPro; IPR007110; IG-1-like.
 DR InterPro; IPR003598; IG_c2.
 DR Pfam; PF00047; Ig_3.
 DR SMART; SM00408; IGC2; 2.
 DR PROSITE; PS50835; IG_LIKE; 3.
 KW Immunoglobulin domain; Cell adhesion; Glycoprotein; Lipoprotein;
 KW Membrane; GPI-anchor; Repeat; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 314
 FT OPIOID BINDING PROTEIN/CELL ADHESION
 FT MOLECULE HOMOLOG.
 FT REMOVED IN MATURE FORM (POTENTIAL).
 FT DOMAIN 32 119
 FT DOMAIN 129 211
 FT DOMAIN 215 302
 FT DISULFID 50 108
 FT DISULFID 150 194
 FT DISULFID 236 288
 FT LIBID 314 314
 FT (Potential).
 FT GPI-anchor amidated asparagine
 FT (Potential).

FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 337 AA; 36887 MW; BAE7175518565651E CRC64;

Query Match 7.6%; Score 100.5; DB 1; Length 337;
 Best Local Similarity 21.5%; Pred. No. 1.4;
 Matches 56; Conservative 38; Mismatches 103; Indels 63; Gaps 11;

OY 4 GVPRHLLIVQLALLPAATGKNVYLGKGDVYELTCTASQKSIQFHKNSQIKLG 63
 DB 21 GVPRRS-----GDATFPFAMDN--VTVRQESATLCTVDYDR-RVAMNRSITLYAG 71
 OY 64 NQGSFLTKGPKNDRADRSRLMDQGNFPLINKL-----IEDSDYICVE 112
 DB 72 N-----DKMSIDRVVILSNITQVSIKIHNVADVDEGPTGVSQ 111
 OY 113 ----DQKEVQLVFG----LTASDTHLLQGQSLTTLSPSSPSVQCRSPRKN-- 162
 DB 112 TDNHPKTSRVHLIVQVPPQIVNISDLTVNNGSSVTLMCLAFGRPEPTVTRHLSGKQG 171
 OY 163 -TQGGKTLVSQLELDQSGTWTCTVLQ----NKKVQFKTIVPRASALPAPPGSALP 216
 DB 172 FVSEDEYLETITGIREQSGEYCSAVNDVAVDVRYKVTNVYPYIS--NAKXTGASVG 229
 OY 217 DPQTRASLPDPAPASALPAA 236
 DB 230 QKGILQC-----EASAVPVA 244

RESULT 74
 CEK2_CHICK STANDARD; PRT; 806 AA.

AC P18460;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tyrosine kinase receptor CEK2 precursor (EC 2.7.1.112).
 GN CEK2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90332672; PubMed=2165604;
 RA Paquale E.B.;
 RT "A distinctive family of embryonic protein-tyrosine kinase
 RT receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5812-5816(1990).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Belongs to the fibroblast growth factor receptor
 CC family.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M35195; AAA4664.1; -
 CC PIR; A35963; A35963.
 DR HSSP; P11362; IFGK.
 DR InterPro; IPR007110; IG-1-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR000719; Prot_kinase.

Query Match	Best local similarity	7.6%	Score 100.5	DB 1	Length 806
Matches 59	Conservative 32	Mismatches 85	Indels 107	Gaps 12	
Qy	10 LLLVLQALLLPATOG	---	NRVLGKKDITVELTCTASQKRSI	49	
Db	11 LCLAAVAGALPAARRRGAERSGQAAEYLRSETAFLEELVPG-SGDTIELSCN-TQSSSV	---		68	
Qy	50 QPHW	---	KNSQIKILNQGSFLTKG	72	
Db	69 SVFPEKDIGIAPNSRTHIGQKLKIIVSYDSDGLGYCKPRHSNE--VLNFTVRYVDS	---		126	
Qy	73 PSKLNDRAISRSLWDQGNFPLIKNLKIEIDSDTYICEVEDQKEVQLLVGELTANSDTH	---		132	
Db	127 PSSGDDDDDDDES	---	EDTGVPFWTRPDKMEKKLLAV--PAAN	164	
Qy	133 LLQGSLLTLLESPPGSSPSVQC	---	RSBRKNIGQKTLVSQLEL	179	
Db	165 -----TVRRCPGAGNPRTPIYVLKNGKEKEGHRIGGIGIKLRHQQSLVMSVSPDR	---		217	
Qy	180 GTMTCTVLAONQKVE--FKRIDVPAASALPAPPTSAALPDPT	---		220	
Db	218 GNYTCVVENKYGKNTIRHTYQDLVLEKSPRRPILQAG--LPAHOT	---		258	

[illegible]

RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
 RA Carriano A.V.;
 RA Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
 RA Satchleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebli T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [5]
 RP MEDLINE=98102450; PubMed=9427723;
 RA Bades-Perner A., Thompson J., van der Putten H., Zimmermann W.;
 RT "Wice transgenic for the human CGM6 gene express its product, the
 RT granulocyte marker CD66b, exclusively in granulocytes";
 RL Blood 91:663-672(1998).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
 CC -1- TISSUE SPECIFICITY: Expressed in leukocytes of chronic myeloid
 CC leukemia patients and bone marrow.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. CEA family.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
 CC -1- DATABASE: NMB=PROW; NOTE=CD guide CD66b entry;
 CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd66b.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; X52378; CAA36604.1; -
 CC EMBL; M33326; AAA59914.1; -
 CC EMBL; D90064; BAA14108.1; -
 CC EMBL; AC004558; AAC13659.1; -
 CC EMBL; BC026263; AAC26263.1; -
 CC EMBL; Z95119; CAB08298.1; -
 CC PIR; S13524; A34815.
 CC Genew; HGNC:1820; CEACAM8.
 CC MIM; 114890; -
 CC GO; GO:0005615; C:extracellular space; TAS.
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0006955; P:immune response; TAS.
 CC InterPro; IPR007110; Ig-like.
 CC Pfam; PF00047; Ig; 3.
 CC PROSITE; PSS0835; IG LIKE; 2.
 CC Immunoglobulin domain; Antigen; Signal; Glycoprotein; GPI-anchor;
 CC Repeat; Polymorphism; Lipoprotein.
 CC SIGNAL 1 34
 CC CHAIN 35 320 CARCINOEMBRYONIC ANTIGEN-RELATED CELL
 CC FT PROPEP 321 349 ADHESION MOLECULE 8.
 CC FT LIPID 320 320 REMOVED IN MATURE FORM.
 CC GPI-anchor amidated aspartate.

FT DOMAIN 35 142 IG-LIKE V-TYPE.
 FT DOMAIN 145 232 IG-LIKE C2-TYPE 1.
 FT DOMAIN 237 319 IG-LIKE C2-TYPE 2.
 FT DISULFID 167 215 PROBABLE.
 FT DISULFID 259 299 PROBABLE.
 FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 114 114 R -> K (in dbSNP:1041597).
 FT FTID=VAR 011721.
 FT CONFLICT 322 322 L -> V (IN REF. 2).
 FT SEQUENCE 349 AA; 38154 MW; AACF74DALAC839D8 CRC64;
 Query Match 7.5%; Score 98.5; DB 1; Length 349;
 Best Local Similarity 20.2%; Pred. No. 2.1;
 Matches 60; Conservative 51; Mismatches 87; Indels 99; Gaps 15;
 QY 5 VPRRHLLVQL-----ALPRAIGKNTVY-----GRKDT 36
 DB 13 IPQGLLTPASLFTFWNPPTTAQLTEAVPSNAEKEVLLVHNPDPDRGYWKGET 72
 QY 37 VELTASQKKSIOFWKNSNOIKLNGOSFLTKGSKLNDRAISRSLMDGCPPLIT 96
 DB 73 VD-----ANKRIITG-----YISNOQ--ITROPATNS-----RETIYV--NLSLM 109
 QY 97 KKLKIEDSDTY---ICEVEDQKEV---QLLVF-----GLTANSDFHLLQGSGLTTLTLES 145
 DB 110 RNVTRDTSYTLQVILKMLMSEEVVGFSYHPEFPKPSISNNNSNPVEDKDAVFTCEP 169
 QY 146 PR-----GSSPVQCRSPRGKNIQGGKTLVSQELDSDGTWTCTVLONOKVER 195
 DB 170 ETQNTTYLWVNGQSLPV--SPRLQLSNGNRTLLSTVRDVGYECEI--ON----- 219
 QY 196 KIDIVPRASALPAPPTGASLPDPQTASALPDPASALPALAVALISFLGLGIVAC 252
 DB 220 -----PASANFSDPYTLNVLYGDPATISPSD---TYHAGVNLNLSC 259
 RESULT 77
 SLIB_HUMAN STANDARD; PRT; 686 AA.
 AC Q96R66;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Static acid binding Ig-like lectin 11 precursor (Siglec-11) (Static
 GN acid-binding lectin 11) (UNQ9222/PRO28718).
 OS SIGLEC11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH PTPN6 AND PTPN11.
 RX MEDLINE=22086217; PubMed=11986377;
 RA Angata T., Kerr S.C., Greaves D.R., Varki N.M., Crocker P.R.,
 RA Varki A.;
 RT "Cloning and characterization of human Siglec-11. A recently evolved
 RT signaling that can interact with SHP-1 and SHP-2 and is expressed by
 RT tissue macrophages, including brain microglia";
 RL J. Biol. Chem. 277:24466-24474(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=22887296; PubMed=12975309;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,

RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Haen D., Foster J., Grimaldi C., Gu O., Haas P.E., Heldens S.,
 RA Huang A., Kim H.S., Kilmowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark W., Robble E., Sanchez C., Schoenfeld J.,
 RA Seasholtz S., Simons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wiand D., Woods K., Xie M.-H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P.;
 RA "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270(2003).
 CC -1- FUNCTION: Putative adhesion molecule that mediates stialic-acid
 CC dependent binding to cells. Preferentially binds to alpha2,8-
 CC linked stialic acid. The stialic acid recognition site may be masked
 CC by cis interactions with stialic acids on the same cell surface. In
 CC the immune response, may act as an inhibitory receptor upon ligand
 CC induced tyrosine phosphorylation by recruiting cytoplasmic
 CC phosphatase(s) via their SH2 domain(s) that block signal
 CC transduction through dephosphorylation of signaling molecules.
 CC -1- SUBUNIT: Interacts with PTPN6/SHP-1 and PTPN11/SHP-2 upon
 CC phosphorylation.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing: Named isoforms=2;
 CC Name=1;
 CC IsoId=Q96RL6-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q96RL6-2; Sequence=VSP 008764;
 CC -1- TISSUE SPECIFICITY: Expressed by macrophages in various tissues
 CC including Kupffer cells. Also found in brain microglia.
 CC -1- DOMAIN: Contains 1 copy of a cytoplasmic motif that is referred
 CC to as the immunoreceptor tyrosine-based inhibitor motif (ITIM).
 CC This motif is involved in downmodulation of cellular responses.
 CC The phosphorylated ITIM motif binds to the SH2 domain of
 CC PTPN6/SHP-1.
 CC -1- PTM: Phosphorylated on tyrosine residues.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
 CC (sialic acid binding Ig-like lectin) family.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, AF373918; AAK72907.1; -;
 DR EMBL, AY358135; AAQ88502.1; -;
 DR Genew; HGNC:15622; SIGLEC11.
 DR MIM; 607157; -;
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_C2.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00408; IGC2; 2.
 DR PROSITE; PS00835; IG LIKE; 3.
 DR PROSITE; PS00290; IG MHC; 1.
 KM Cell adhesion; Repeat; Signal; Transmembrane; Immunoglobulin domain;
 RT Lectin; Glycoprotein; Phosphorylation; Alternative splicing.
 RT SIGNAL
 FT SIGNAL 1 15
 FT CHAIN 16 686
 FT DOMAIN 16 549
 FT TRANSMEM 550 572
 FT SIGNAL 573 686
 FT DOMAIN 19 122
 FT DOMAIN 147 232
 FT DOMAIN 239 338
 FT DOMAIN 343 440
 FT SIGNAL 630 635
 FT ITIM MOTIF.

FT DISULFID 37 174 BY SIMILARITY.
 FT DISULFID 42 102 BY SIMILARITY.
 FT DISULFID 165 216 BY SIMILARITY.
 FT DISULFID 275 322 BY SIMILARITY.
 FT DISULFID 379 424 BY SIMILARITY.
 FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 78 78 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 443 539 YPPOLLGPGSCVBAAGHSCSSQSPAPSLRMWLGELLE
 FT GNSGSGSPREYTPSAGPWANSSLSLHGSLSLRCAMN
 FT VHGQSGSVFOLLG -> W (in isoform 2).
 FT FTId=VSP_008764.
 FT CONFLICT 84 84 E -> A (in REF. 2).
 FT CONFLICT 353 353 A -> G (in REF. 2).
 SQ SEQUENCE 686 AA; 74544 MW; F00EA1936156E5A8 CRC64;
 Query Match Best Local Similarity 7.5%; Score 98.5; DB 1; Length 686;
 Matches 61; Conservatly 20.7%; Pred. No. 4.7;
 Mismatches 99; Indels 97; Gaps 12;
 QY 24 QGNRVVIG-KKGDVVELCTASOKKSIQPHKNSNQKIIGNQSPFLTKPSKLNDRADS 82
 DB 257 QGNVYVLEVGQGRFLRLCAADSQPPATLSH-----VLQDRVLS 295
 QY 83 RRLMDQGNPFLIKNLKIEDSDTYICEVEDQ-----KEEVOLLVF----- 123
 DB 296 SSHWNGRTGLGLRGVAGSGYTRAEARLRLSQQAADLSVQYPENLRVWVSQNR 355
 QY 124 ----GLTNSDTHLLQGOSLT---TLSPGSSPSVQCSPPRKNTQGGKTLSSVQ--- 173
 DB 356 TVLENLNGTSLPTLEGOSLRVLCVTHSSPPA-----RLSWTRMGQTVGPSQPSD 405
 QY 174 -----LELQDSGWTCTV---LQNRK-KVEFKDIVPR-----A 203
 DB 406 PGVLELPPIQWHEGEFTCAQHPLGSOHVSLSVHYPPOLLGPPSCSWAEGHCSSS 465
 QY 204 SALPAPP-----TGSALPDPQTASALPDPASALPALAVLSFLLGL--GLGVAC 252
 DB 466 QASAPAPLRMWLGSELLEGNSQGSFVTPSSAPPMANSSLSLHGSSGLRLNC 520
 RESULT 78
 CD8B_SAISC STANDARD; PRT; 209 AA.
 ID CD8B_SAISC
 AC Q9XSM7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE T-cell surface glycoprotein CD8 beta chain precursor (Antigen CD8B).
 GN CD8B.
 OS Saimiri sciureus (Common squirrel monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
 OX NCBI_Taxid=9521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SMPAIN=92039;
 RX MEDLINE=99299362; PubMed=10369934;
 RA Ureca-Vidal A., Garcia Z., Lemoultier F.A., Kazanjli M.;
 RT "Molecular characterization of cDNAs encoding squirrel monkey (Saimiri
 RL sciureus) CD8 alpha and beta chains.";
 CC Immunogenetics 49:718-721(1999).
 CC -1- FUNCTION: Identifies cytotoxic/suppressor T-cells that interact
 CC with MHC class I bearing targets. CD8 is thought to play a role in
 CC the process of T-cell mediated killing.
 CC -1- SUBUNIT: In general heterodimer of an alpha and a beta chain
 CC linked by two disulfide bonds.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ130819; CAB1463.1;
DR GO: GO:0042101; C-T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042288; F:MHC class I protein binding; ISS.
DR GO: GO:0005515; F:protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0042110; P:T-cell activation; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; Ig; 1.
DR POSITE: PS50835; IG LIKE; 1.
KM Immunoglobulin domain; Transmembrane; T-cell; Antigen; Glycoprotein;
KW Immune response; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 209 T-CELL SURFACE GLYCOPROTEIN CD8 BETA
FT CHAIN 22 209 CHAIN.
FT DOMAIN 22 169 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 170 190 POTENTIAL.
FT DOMAIN 191 209 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 22 122 IG-LIKE V-TYPE.
FT DISULFID 41 115 POTENTIAL.
FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 209 AA; 23481 MW; E595E8E39DF629C5 CRC64;

Query March 7.4%; Score 98; DB 1; Length 209;
Best Local Similarity 26.1%; Pred. No. 1.2;
Matches 46; Conservative 27; Mismatches 73; Indels 30; Gaps 8;

QY 6 PFRHLVLVLQALPAATQGNKVV-----VLAKKGDTVELTCTASQKSIQFHM----- 53
DB 3 PRMWLLSAQL-----ALHNGSVLQGPAYIMVQTNQMWLSCAIISSSTRIYVQLH 58
QY 54 --KNSNOIKILGNOSFLTKPSKLNDRADSRRLMOCN-FPLITKLIEDSDTYICE 110
DB 59 APSNSNHEILAFWDS--SKGTIHSSEGVOKKITVFRDGLFFPLNLRVXLEDSGYFCM 116
QY 111 VEDQKEVQLVLPGLTANSPTHLIQGSLTLTLESPPGSSP-SVQGRSPGKNIOG 165
DB 117 VIGS-----PTLIFG-----TGTQLSVVDILPTTAQTKKSTPKVTVCRLPRPRTKRG 164

RESULT 79
PIGR_HUMAN STANDARD; PRT; 764 AA.
AC P01833;
DC 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Polymetric-immunoglobulin receptor precursor (poly-Ig receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=92039621; PubMed=1682231;
RA Kraljci P., Grzeschik K.H., Geurts van Kessel A.H., Olausen B.,
RA Brandtzaeg P.;
RT "The human transmembrane secretory component (poly-Ig receptor):
RT molecular cloning, restriction fragment length polymorphism and

```

```

RT Chromosomal sublocalization.";
RL Hum. Genet. 87:642-648 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92387236; PubMed=1355431;
RA Kraljci P., Kvale D., Taeken K., Brandtzaeg P.;
RT "Molecular cloning and exon-intron mapping of the gene encoding human
RT transmembrane secretory component (the poly-Ig receptor).";
RL Eur. J. Immunol. 22:2309-2315 (1992).
RN [3]
RP SEQUENCE OF 72-764 FROM N.A.
RX MEDLINE=89149795; PubMed=2920039;
RA Kraljci P., Solberg R., Sandberg M., Oyen O., Janssen T.,
RA Brandtzaeg P.;
RT "Molecular cloning of the human transmembrane secretory component
RT (poly-Ig receptor) and its mRNA expression in human tissues";
RL Biochem. Biophys. Res. Commun. 158:783-789 (1989).
RN [4]
RP SEQUENCE OF 19-577, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=85128981; PubMed=6526384;
RA Biffert H., Quentin E., Decker J., Hillemeir S., Hufschmidt M.,
RA Klingmuller D., Weber M.H., Hilschmann N.;
RT "The primary structure of human free secretory component and the
RT arrangement of disulfide bonds.";
RL Hoppe-Seyler's Z. Physiol. Chem. 365:1489-1495 (1984).
RN [5]
RP SEQUENCE OF 19-577.
RX MEDLINE=91315750; PubMed=1859628;
RA Biffert H., Quentin E., Decker J., Hillemeir S., Decker J.,
RA Weber M., Hilschmann N.;
RT "Determination of the molecular structure of the human free secretory
RT component.";
RL Biol. Chem. Hoppe-Seyler 372:119-128 (1991).
RN [6]
RP SEQUENCE OF 118-138; 212-230; 232-268; 273-288 AND 578-603.
RX MEDLINE=97379357; PubMed=9237679;
RA Hughes G.J., Frutiger S., Savoy L.-A., Reason A.J., Morris H.R.,
RA Jaton J.-C.;
RT "Human free secretory component is composed of the first 585 amino
RT acid residues of the polymetric immunoglobulin receptor.";
RL FEBS Lett. 410:443-446 (1997).
CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Also secreted.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: S62403; AAB20203.1; -
DR EMBL: S43449; AAB23176.1; -
DR EMBL: S43437; AAB23176.1; JOINED.
DR EMBL: S43441; AAB23176.1; JOINED.
DR EMBL: S43442; AAB23176.1; JOINED.
DR EMBL: S43443; AAB23176.1; JOINED.
DR EMBL: S43444; AAB23176.1; JOINED.
DR EMBL: S43445; AAB23176.1; JOINED.
DR EMBL: S43446; AAB23176.1; JOINED.
DR EMBL: S43447; AAB23176.1; JOINED.
DR EMBL: S43448; AAB23176.1; JOINED.
DR EMBL: M24559; AAA36102.1; -
DR EMBL: A52091; CAA03384.1; -
DR PIR: A46537; ORHUGS.
DR GlycoStatedB: P01833; -

```


DR Genew; HGNC:8968; PIGR.
 DR MIM:173880; -
 DR GO:GO:0005887; C:integral to plasma membrane; TAs.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR Pfam; PF00047; Ig; 5.
 DR SMART; SM00409; IG; 5.
 DR PROSITE; PS00835; IG-LIKE; 2.
 DR Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
 KW Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 764
 FT CHAIN 19 603
 FT DOMAIN 19 638
 FT TRANSMEM 639 661
 FT DOMAIN 662 764
 FT DOMAIN 19 120
 FT DOMAIN 145 237
 FT DOMAIN 250 352
 FT DOMAIN 364 458
 FT DOMAIN 462 561
 FT DISULFID 40 110
 FT DISULFID 56 64
 FT DISULFID 152 220
 FT DISULFID 257 325
 FT DISULFID 271 279
 FT DISULFID 371 441
 FT DISULFID 385 395
 FT DISULFID 482 544
 FT DISULFID 486 520
 FT DISULFID 496 503
 FT CARBOHYD 83 83
 FT CARBOHYD 90 90
 FT CARBOHYD 135 135
 FT CARBOHYD 186 186
 FT CARBOHYD 421 421
 FT CARBOHYD 469 469
 FT CARBOHYD 499 499
 FT VARIANT 580 580
 FT CONFLICT 136 136
 FT CONFLICT 158 158
 FT CONFLICT 208 209
 FT CONFLICT 229 229
 FT CONFLICT 234 234
 FT CONFLICT 241 241
 FT CONFLICT 262 262
 FT CONFLICT 280 280
 FT CONFLICT 392 392
 FT CONFLICT 500 500
 SQ SEQUENCE 764 AA; 83313 MW; 916B3B62C339950 CRC64;
 Query Match 7.4%; Score 98; DB 1; Length 764;
 Best Local Similarity 20.2%; Pred. No. 5.9;
 Matches 53; Conservative 36; Mismatches 101; Indels 72; Gaps 10;
 Oy 2 NRGVPRHLLVLTALPAATQGNKVLGKKDTELTCT---TASQKSIQFHWKNS 56
 Db 116 SRGSLFD---VSLVSGPGLNDTQKTVVLTGRTVINCPTENAKRKSL----- 165
 Oy 57 NQITLNGQGSFLTKGPKSLMDRADSRSLMDQGN---FLITKNLKIETSDTYICEVE 112
 Db 166 --YQIGLYPVLVDDSGYVNPNTGRIRLDIQGGQLLFVSVINQRLSDAGQYLCQAG 223
 Oy 113 DQKEVQLVFGTLTANSPTHLLOGSLTLTLESPGS-----SPSVQ-----CRSPR 159
 Db 224 DDSNKN-----KNADLVLKPEP-ELVYEDLRGSGVTFHCAIGPEVANYAKFLCQSS 275
 Oy 160 GKNI-----GCKTILSVQLELDQSGTWCTTV-----LQN 189
 Db 276 GENDVVVNTLGRKAPAFEGRIILNPDQKDSFVSVITGLRKEKDGARYLCAHSDGQLOE 335
 Oy 190 QKVFEDKIDVPRASALPAPT 211

Db 336 GSPIQAMQLFVNEESTIPRSPT 357
 RESULT 80
 ID TYO3 MOUSE STANDARD; PRT; 880 AA.
 AC P51144;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tyrosine-protein kinase receptor TYRO3 precursor (EC 2.7.1.112)
 DE (Tyrosine-protein kinase RSE) (Tyrosine-protein kinase DTK) (TK19-2).
 GN TYRO3 OR DTK OR RSE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94193774; PubMed=7511603;
 RA Mark M.R., Scadden D.T., Wang Z., Gu Q., Goddard A., Godowski P.J.;
 RT "RSE, a novel receptor-type tyrosine kinase with homology to Axl/Ufo,
 RL is expressed at high levels in the brain.";
 RL J. Biol. Chem. 269:10720-10728 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95161079; PubMed=7857657;
 RA Crostier P.S., Lewis P.M., Hall L.R., Vitas M.R., Morris C.M.,
 RA Beier D.R., Wood C.R., Crostier K.E.;
 RT "Isolation of a receptor tyrosine kinase (DTK) from embryonic stem
 RL cells: structure, genetic mapping and analysis of expression.";
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=94336210; PubMed=8058320;
 RA Lai C., Gore M., Lemke G.;
 RT "Structure, expression, and activity of Tyro 3, a neural adhesion-
 RL related receptor tyrosine kinase.";
 RN [4]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=95240399; PubMed=7723626;
 RA Schulz N., Pauliac C., Lee L., Zhou R.;
 RT "Isolation and expression analysis of tyro3, a murine growth factor
 RL receptor tyrosine kinase preferentially expressed in adult brain.";
 RL Brain Res. Mol. Brain Res. 28:273-280 (1995).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX STRAIN=BALB/c; TISSUE=Brain;
 RA Sasaki M.;
 RT Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May be involved in cell adhesion processes, particularly
 CC in the central nervous system.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Abundant in the brain and lower levels in
 CC other tissues.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
 CC AXL/UFO SUBFAMILY.
 CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U05683; AAA19237.1; -

DR EMBL; U18933; AAC52148.1; -

DR EMBL; X78103; CAAS4995.1; ALT INIT.

DR EMBL; U18342; AAB26942.1; ALT_INIT.

DR EMBL; AB000828; BAA19193.1; -

DR PIR; B53743; B53743.

DR HSSP; P11362; IFGK.

DR MGJ; MG1.104294; TYR3.

DR InterPro; IPR008957; FN_III-like.

DR InterPro; IPR003961; FN_III.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR001245; Tyr_kinase.

DR InterPro; IPR008266; Tyr_kinase_AS.

DR Pfam; PF00041; fn3; 2.

DR Pfam; PF00047; Ig; 2.

DR Pfam; PF00069; kinase; 1.

DR PRINTS; PR00109; TYRKINASE.

DR PRODOM; PD000001; Prot_kinase; 1.

DR SMART; SM00408; IGc2; 1.

DR SMART; SM00219; Tyrc; 1.

DR PROSITE; PS00835; IG-LIKE; 2.

DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.

DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN KINASE_TYR; 1.

KW Cell adhesion; Transferrase; Tyrosine-protein kinase; Receptor;

KW ATP-binding; Repeat; Signal; Transmembrane; Immunoglobulin domain;

KW Glycoprotein; Phosphorylation.

FT SIGNAL 1 30

FT CHAIN 31 880

FT DOMAIN 31 419

FT TRANSMEM 420 440

FT DOMAIN 441 880

FT DOMAIN 31 118

FT DOMAIN 129 209

FT DOMAIN 214 301

FT DOMAIN 305 401

FT DOMAIN 508 785

FT NP_BIND 514 522

FT BINDING 540 540

FT ACT_SITE 645 645

FT DISULFID 54 107

FT DISULFID 150 193

FT CARBOHYD 53 53

FT CARBOHYD 75 75

FT CARBOHYD 181 181

FT CARBOHYD 220 220

FT CARBOHYD 230 230

FT CARBOHYD 283 283

FT CARBOHYD 356 356

FT CARBOHYD 370 370

FT MOD_RES 676 676

FT CONFLICT 630 630

FT CONFLICT 811 811

SEQ SEQUENCE 880 AA; 96222 MW; 387AC36CB016B3F7 CRC64;

Query Match 7.4%; Score 98; DB 1; Length 880;

Best Local Similarity 20.1%; Pred. No. 7;

Matches 59; Conservative 44; Mismatches 97; Indels 94; Gaps 11;

QY 4 GVPRHLLVQLA--LLPATQGNKVLG-----KSGDVELTCTASQKSIQPHW- 53

DB 8 GWPGLRPLLAGLALLPGSAAGLKLMAFPVMTYSGQGPVTLNCSVEGMEPDTHM 67

QY 54 -----KSNQIKILGNQGSFLTKGPKLNDRAISRSLMDQGNFLLIKLKIEDSDTY 107

DB 68 KDTGVVONASQVSISIEHSGWGL-----LSLKSVERSDAGLY 105

QY 108 ICEVEDQKEEVOLVGLTANSPTHLLOQSLLTLES-----PGSSPSVQ 154

DB 106 WCVVND-GEETKI-----SQSWLVTEGVFPFTVEPKDLAVPNAPFQIS 149

QY 155 CRSP-----RGKNIQGGKTLVSQLE---LQDSGTWCTVLAQNKVKEFKIDIV 200

DB 150 CEAVGPPEPVITYMMGLTKVGSPAPSVLVNTGVTQGTERTSCSA-RNKKGLATSRPAI 208

QY 201 PRASALPAPPTGSALPDPTAGALPDPPASALPALAVISFLGLGLGVACVL 254

DB 209 VRLQAPPAPAFNTV-----TTISSYNASVAVWPGADGIALHSCTV 250

RESULT 81

ID KVST_MOUSE STANDARD; PRT; 108 AA.

AC P01653;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig kappa chain V-V region W3082.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RN [1]

RP SEQUENCE.

RX MEDLINE=82099361; PubMed=6798111;

RA Johnson N., Slinkard J., Paul L., Hood L.;

RT "The complete V domain amino acid sequences of two myeloma inulin-binding proteins";

RT J. Immunol. 128:302-307(1982).

CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).

DR PIR; B2811; KYMS82.

DR HSSP; P80362; IWTL.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; Ig; 1.

DR SMART; SM00406; Icy; 1.

KW PROSITE; PS00835; IG-LIKE; 1.

KW Immunoglobulin V region.

FT DOMAIN 1 23

FT DOMAIN 24 34

FT DOMAIN 35 49

FT DOMAIN 50 56

FT DOMAIN 57 88

FT DOMAIN 89 97

FT DOMAIN 98 108

FT DISULFID 23 88

FT NON_TER 108 108

SEQ SEQUENCE 108 AA; 11850 MW; C5C145DC376F30CD CRC64;

Query Match 7.4%; Score 97.5; DB 1; Length 108;

Best Local Similarity 37.7%; Pred. No. 0.58;

Matches 29; Conservative 6; Mismatches 37; Indels 5; Gaps 2;

QY 34 GDTVELTCTASQKSIQFHKNSNQIKILGNQGSFLTKGPKLNDRAISRSLMDQ-NF 92

DB 16 GDIVTWTCAQSQGTININLMFOQKP-----GKAPKLIVGASILEDGVPSRFGSGRYGTF 71

DB 72 TLTISLEDDMATYFC 88

RESULT 82

ID ICOL_MOUSE STANDARD; PRT; 322 AA.

AC Q9JH8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE ICOS ligand precursor (B7 homolog 2) (B7-H2) (B7-like protein G150)
DE (B7-related protein-1) (B7RP-1) (LICOS).
CN ICOSL OR B7H2 OR B7RP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN
RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RC TISSUE=Lymphocytes;
RX MEDLINE=20083495; Pubmed=10617205;
RA Yoshinaga S.K., Whoriskey J.S., Khare S.D., Samiento U., Guo J.,
R Horan T., Shih G., Zhang M., Cocci M.A., Kohno T., Tafuri-Bladt A.,
R Brankow D., Campbell P., Chang D., Chi L., Dai T., Duncan G.,
RA Elliott G.S., Hui A., McCabe S.M., Scully S., Shahinian A.,
R Shaklee C.L., Van G., Mak T.W., Senaldi G.;
RT "T-cell co-stimulation through B7RP-1 and ICOS".
RL Nature 402:827-832(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Thymus;
RX MEDLINE=20015817; Pubmed=10549624;
RA Swallow M.M., Wallin J.J., Sha W.C.;
RT "B7h, a novel costimulatory homolog of B7.1 and B7.2, is induced by
RT TNFalpha.";
RL Immunol 11:423-432(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C3H/HeJ; TISSUE=Fetal thymus;
RX MEDLINE=20126021; Pubmed=10557666;
RA Ling V., Wu P.W., Finerty H.F., Bean K.M., Spaulding V., Fouser L.A.,
R Leonard J.P., Hunter S.E., Zoller R., Thomas J.L., Miyashiro J.S.,
RA Jacobs K.A., Collins M.;
RT "Identification of GL50, a novel B7-like protein that functionally
RT binds to ICOS receptor.";
RL J. Immunol. 164:1653-1657(2000).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=PerIPHERAL blood lymphocytes;
RX MEDLINE=21286479; Pubmed=11390480;
RA Ling V., Wu P.W., Miyashiro J.S., Marusic S., Finerty H.F.,
RA Collins M.;
RT "Differential expression of inducible costimulator-ligand splice
RT variants: lymphoid regulation of mouse g150-b and human g150
RT molecules.";
RL J. Immunol. 166:7300-7308(2001).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA Ling V., Dunnssel-Dannopoloulos K.;
RT "G150 molecules and uses thereof.";
RL Patent number WO0121796, 29-MAR-2001.
CC -I- FUNCTION: LIGAND FOR THE T-CELL-SPECIFIC CELL SURFACE RECEPTOR
CC ICOS. ACTS AS A COSTIMULATORY SIGNAL FOR T-CELL PROLIFERATION AND
CC CYTOKINE SECRETION; INDUCES ALSO B-CELL DIFFERENTIATION AND
CC DIFFERENTIATION INTO PLASMA CELLS. COULD PLAY AN IMPORTANT ROLE IN
CC MEDIATING LOCAL TISSUE RESPONSES TO INFLAMMATORY CONDITIONS, AS
CC WELL AS IN MODULATING THE SECONDARY IMMUNE RESPONSE BY CO-
CC STIMULATING MEMORY T CELL FUNCTION. DURING PREGNANCY, MAY FUNCTION
CC TO skew THE CYTOKINE OF MATERNAL T-CELLS TOWARD IMMUNOPROTECTIVE
CC THIS PHENOTYPE.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=1;
CC IsoId=O9UH8-1, Sequence=Displayed;
CC Name=2; Synonyms=B;
CC IsoId=O9UH8-2; Sequence=VSP_002521;
CC -I- TISSUE SPECIFICITY: ISOFORM 1 HIGHEST EXPRESSION IN LYMPHOID
CC TISSUES, SUCH AS SPLEEN (MOSTLY IN THE MARGINAL ZONE), LYMPH NODES
CC (PARTICULARLY IN THE CORTEX AND IN BOTH PRIMARY AND SECONDARY
CC FOLLICLES), THYMUS (PREDOMINANTLY IN THE MEDULLA) AND PEYER'S
CC PATCHES (MOSTLY IN THE FOLLICLES), LOWER LEVELS IN MANY

CC		NONLIMPHOID TISSUES, SUCH AS BRAIN, HEART, KIDNEY, LIVER, LUNG,
CC		SKELETAL MUSCLE AND TESTIS. PRESENT ON FRESHLY ISOLATED SPLENIC B-
CC		CELLS, T-CELLS, DENDRITIC CELLS AND MACROPHAGES. THE EXPRESSION OF
CC		ISOFORM 2 IS RESTRICTED TO HEART, SPLEEN AND KIDNEY.
CC	-1	DEVELOPMENTAL STAGE: DETECTED EARLY IN HEMOPOIESIS; IN THE YOLK
CC		SAC AT 11.5 AND 12.5 DPC AND, TO A LESSER EXTENT, IN THE LAYER AT
CC		14.5 DPC.
CC	-1	SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
CC		family.
CC	-1	SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC	-1	SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (see http://www.isb-stb.ch/announce/
CC		or send an email to license@isb-stb.ch).
CC		-----
DR	EMBL, AF216747, AAFA5149.1;	-
DR	EMBL, AF199027, AAFA4738.1;	-
DR	EMBL, AX100591, CAC36463.1;	-
DR	EMBL, AX100593, CAC36464.1;	-
DR	EMBL, AF394451, AAK77544.1;	-
DR	MGI, MGI:1354701, ICob1.	
DR	GO, GO:0005576, C:extracellular; IDA.	
DR	GO, GO:0016021, C:integral to membrane; NAS.	
DR	GO, GO:0003793, F:defense/immunity protein activity; NAS.	
DR	GO, GO:0005102, F:receptor binding; TAS.	
DR	GO, GO:0045190, F:isotype switching; NAS.	
DR	GO, GO:0042104, P:positive regulation of activated T-cell pro. . ; TAS.	
DR	GO, GO:0045404, P:positive regulation of interleukin-4 biosyn. . ; IMP.	
DR	GO, GO:0007165, P:signal transduction; NAS.	
DR	GO, GO:0042110, P:T-cell activation; NAS.	
DR	InterPro, IPR007110, Ig-like.	
DR	InterPro, IPR003599, Ig.	
DR	Pfam, PF00047, Ig; 1.	
DR	SMART, SM00409, IG; 1.	
DR	PROSITE, PS50835, IG LIKE; 2.	
KM	B-cell activation; Immune response; Glycoprotein;	
KM	Immunoglobulin domain; Signal; Transmembrane; Multigene family;	
KM	Alternative splicing.	
FT	SIGNAL	1..46
FT	CHAIN	47..322
FT	DOMAIN	47..277
FT	TRANSHEM	278..298
FT	DOMAIN	299..322
FT	DOMAIN	49..149
FT	DOMAIN	167..263
FT	DOMAIN	31..38
FT	DOMAIN	289..292
FT	DISULFID	62..138
FT	DISULFID	185..243
FT	CARBOHYD	71..71
FT	CARBOHYD	120..120
FT	CARBOHYD	163..163
FT	CARBOHYD	200..200
FT	CARBOHYD	213..213
FT	CARBOHYD	252..252
FT	CARBOHYD	265..265
FT	VARSPLIC	321..322
FT		/ftid=vsp_002521.
FT		R -> H (IN REF. 4 AND 5; CAC36464).
FT	SEQUENCE	322 AA; 35960 MW; 55CBA4AD12B47B6 CRC64;
Query Match		7.4%; Score 97.5; DB 1; Length 322;
Best Local Similarity		21.5%; Pred. No. 2.2;
Matches		63; Conservative 39; Mismatches 104; Indels 87; Gaps 13;
Yy	10 LLVLTQLALLPAATQGNKVVLGGKKGVDELICTASQK-----SIQFMKNNSN-QIKI--61	

DB 35 LLLSLSCASAEETEVGAVY-----GSNVVLSCIDPHRRHNLGLVYVWQIENPEVSVTY 90
 QY 62 -----LGNQSFLLTKGSPKLANDRADSRRLMDQGNFLIITKLIKEDSDTYICE--- 110
 DB 91 YLPKSPGIVNDSSYKNGRHLST-----DSMK-----QGNFSLYKKNVLPDPTOEFTRVFM 142
 QY 111 -----VEDQKEEVQLLVFGILTANSDPHLLQGGSLTFTLSPGSSPSVQCRPRG---K 161
 DB 143 NTATELVKILIEEYVRLVY---AANFSTPIVISTD-----SSNFGQERTTYCMKNGYPER 194
 QY 162 NIOGKTLVSQLELODGSWTCTVLQNKQKVEFKIDIVPRASALPAP----- 209
 DB 195 NLVYINTTDSLID-----TALQNNVTYINKGLYDVISTLRLEPMTSRGVDLCVE 245
 QY 210 -----PTGSALPDPOQA---SALPDPPAASALPALAVISFLL 244
 DB 246 NVALHONITSISOAESFTGNNTKNPDETHNNELKVLVPLVALVAAAFVSFI 298

RESULT 83
 SMP_COTUA ID SMP_COTUA STANDARD; PRT; 620 AA.
 AC Q92154;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Schwann cell myelin protein precursor (Siglec-4b).
 GN SMP.
 OS Ccuturnix ccuturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Ccuturnix.
 NCBI_Taxid=93934;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-28; 120-132; 135-157 AND 563-571.
 RX MEDLINE=92153423; PubMed=1739462;
 RA Dulac C., Tropak M.B., Cameron-Curry P., Rossier J., Marshak D.R., Roder J., le Douarin N.M.;
 RT "Molecular characterization of the Schwann cell myelin protein, SMP: structural similarities within the immunoglobulin superfamily."; Neuron 8:323-334 (1992).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Exclusively expressed by myelinating and nonmyelinating schwann cells and oligodendrocytes.
 CC -1- DEVELOPMENTAL STAGE: First synthesized at embryonic day 5, it remains expressed by cultured Schwann cells.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC (sialic acid binding Ig-like lectin) family.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL; S83711; AAB21466.1; -.
 DR PIR; JH0593; JH0593.
 DR HSSP; P56276; ITLK.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_C2.
 DR Pfam; PF00047; Ig_2.
 DR SMART; SM00408; IgC2; 2.
 DR PROSITE; PSS0835; IG_LIKE; 2.
 KW Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
 KW Immunoglobulin domain; Repeat.
 FT SIGNAL 1 17
 FT CHAIN 18 620 SCHWANN CELL MYELIN PROTEIN.
 FT DOMAIN 18 516 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 517 536 POTENTIAL.
 FT DOMAIN 537 620 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 28 106 IG-LIKE V-TYPE.
 FT DOMAIN 151 233 IG-LIKE C2-TYPE 1.
 FT DOMAIN 239 322 IG-LIKE C2-TYPE 2.
 FT DOMAIN 325 407 IG-LIKE C2-TYPE 3.
 FT DOMAIN 414 495 IG-LIKE C2-TYPE 4.
 FT DISULFID 35 164 BY SIMILARITY.
 FT DISULFID 40 99 BY SIMILARITY.
 FT DISULFID 158 216 BY SIMILARITY.
 FT DISULFID 260 304 BY SIMILARITY.
 FT DISULFID 346 391 BY SIMILARITY.
 FT DISULFID 420 429 BY SIMILARITY.
 FT DISULFID 431 488 BY SIMILARITY.
 FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 620 AA; 66943 MW; 004B3ECTEDC18FBA CRC64;

Query Match 7.4%; Score 97.5; DB 1; Length 620;
 Best Local Similarity 20.3%; Pred. No. 5;
 Matches 48; Conservative 36; Mismatches 91; Indels 61; Gaps 10;

QY 33 KGDIVELTCTASQKSIQFHMKNSNOIKILNGQSFLLTKGSPKLANDRADSRRLMDQGNF 92
 DB 252 EGSDELGCEAGRPAPPLISWFRGSEV-----LREBPG-----NL 287
 QY 93 PLIKLIKIEDSDTYICEVEDQ-----KEVQVLVFG-----GLTANSDPHLLQGG--SLTYLT 142
 DB 288 RLILSVGPDGDSFECVAVENRHRNRSLOLRVAVAPAPVINGSLWVSGDPVSVTR 347
 QY 143 LESPPGSSPSVQCRSPRGKNGIOGK-----TISVQLELODGSWTCTVQLQNG 190
 DB 348 AESEPAALITV-----LRGKVMAAAIYEDHTVMTMRPARPEDGGTYSC-VAENQH 397
 QY 191 --KXVEFKIDIVPRASALPAP---PTGSALPDPOQTASALPDPPAASALPALAVIS 241
 DB 398 GASSTSFNISVEYPLVLPASRCTAGGAGVRCVCMVNSIPDSSLVPELPTRNQTVS 453

RESULT 84
 PIGR_MOUSE ID PIGR_MOUSE STANDARD; PRT; 771 AA.
 AC Q70570;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE PolymERIC-Immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR) [contains: Secretory component].
 GN PIGR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=95138517; PubMed=7836758;
 RA Piskurich J.F., Blanchard M.H., Youngman K.R., France J.A., Kaetzel C.S.;
 RT "Molecular cloning of the mouse polymERIC Ig receptor. Functional regions of the molecule are conserved among five mammalian species."; J. Immunol. 154:1735-1747 (1995).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvJ;
 RX MEDLINE=98072444; PubMed=9409786;
 RA Martin M.G., Gutierrez E.M., Lam J.T., Li T.W.H., Wang J.;
 RT "Genomic cloning and structural analysis of the murine polymERIC receptor (PIGR) gene and promoter region."; Gene 201:189-197 (1997).

```

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Liver;
RX MEDLINE=99410926; PubMed=10481312;
RA de Groot N., Van Kulk-Romeijn P., Lee S.H., de Boer H.A.;
RT "Over-expression of the murine polymorphic immunoglobulin receptor gene
RL in the mammary gland of transgenic mice.";
Transgenic Res. 8:125-135(1999).
CC - FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Type I membrane protein. Also secreted.
CC - SIMILARITY: Contains 5 immunoglobulin-like V-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06431; AAA67440.1; -.
DR EMBL; U83434; AAC53585.1; -.
DR EMBL; U83427; AAC53585.1; JOINED.
DR EMBL; U83428; AAC53585.1; JOINED.
DR EMBL; U83429; AAC53585.1; JOINED.
DR EMBL; U83430; AAC53585.1; JOINED.
DR EMBL; U83431; AAC53585.1; JOINED.
DR EMBL; U83432; AAC53585.1; JOINED.
DR EMBL; U83433; AAC53585.1; JOINED.
DR EMBL; Y16524; CAAT6272.1; -.
DR EMBL; Y16525; CAAT6272.1; JOINED.
DR EMBL; Y16526; CAAT6272.1; JOINED.
DR EMBL; Y16527; CAAT6272.1; JOINED.
DR EMBL; Y16528; CAAT6272.1; JOINED.
DR EMBL; Y16529; CAAT6272.1; JOINED.
DR EMBL; Y16530; CAAT6272.1; JOINED.
DR EMBL; Y16531; CAAT6272.1; JOINED.
DR EMBL; Y16532; CAAT6272.1; JOINED.
DR MGI; MGI:103080; P1gr.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 3.
DR PROSITE; PS50835; IG LIKE; 3.
KM Immunoglobulin domain; Repeat: Transmembrane; Glycoprotein; Signal.
FT CHAIN 1 18
FT CHAIN 19 771
FT CHAIN 19 611
FT CHAIN 19 645
FT TRANSMEM 666 668
FT TRANSMEM 669 771
FT DOMAIN 21 120
FT DOMAIN 135 237
FT DOMAIN 245 351
FT DOMAIN 352 457
FT DOMAIN 463 563
FT DISULFID 40 110
FT DISULFID 152 220
FT DISULFID 257 324
FT DISULFID 370 440
FT DISULFID 484 546
FT CARBOHYD 90 147
FT CARBOHYD 147 170
FT CARBOHYD 206 226
FT CARBOHYD 420 420
FT CARBOHYD 471 471
FT CONFLICT 159 159
A -> V (IN REF. 1).

```

```

FT CONFLICT 396 396 V -> A (IN REF. 1).
FT CONFLICT 620 620 G -> R (IN REF. 1).
SQ SEQUENCE 771 AA; 84998 MW; 78C81302EC710730 CXC64;
Query Match 7.4%; Score 97.5; DB 1; Length 771;
Best Local Similarity 23.0%; Pred. No. 6.5;
Matches 59; Conservative 43; Mismatches 91; Indels 63; Gaps 13;
OY 15 QALLPATQGNK--VYLGRKGDVVELTCTASQKSIQF-----HWK--NSNQIKILND 65
DB 342 QLPVNESSTIINRRSVKGVGVGVAIACPNPKESSLKWKYCKWEGDGNHCPVLVGTQ 401
OY 66 GSFLTKGSKINDADSRSLMDP---GNFPLIKNLKIEDSDPYIC----- 109
DB 402 -----AQVQEEYEGRLALPDGPNGCTYVILNQLTTERAGFYWCLTNDSKRRITTE 453
OY 110 -EVEDQKEEVLVFGLTANSDFHLQGSITLTLESPPG---SSPSVOCR-SPRGKNI-- 163
DB 454 LQVLEATREPLYLEV---TPQWATAVL-GETFTVSCHYPCRFYSGEKWKCKSNKCHILP 509
OY 164 ---QGKTLVS-----QLELDQSGTWCTVLDONKKVEFKDIVRASALP 207
DB 510 SHDEGAROSSVSCDQSSQLVSMTLNPSKEDGVMGCVKQGYGTETAIYI-----AVE 565
OY 208 APPTGSLPDPQTASA 223
DB 566 ERTGSSHVNPFTDANA 581
RESULT 85
VGR1_MOUSE STANDARD; PRT: 1333 AA.
ID VGR1_MOUSE
AC P35969; O55094; Q61517;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vascular endothelial growth factor receptor 1 precursor (EC 2.7.1.112)
DE (VEGFR-1) (Tyrosine-protein kinase receptor FLT) (FLT-1) (Embryonic
DE receptor kinase 2).
GN FLT1 OR FLT OR EPRK2.
OS Mus musculus (Mouse).
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Neonatal brain, and Placenta;
RX MEDLINE=9330572; PubMed=8393164;
RA Finerly H., Kelleher K., Morris G.E., Bean K., Merberg D.M.,
RA Kriz R., Morris J.C., Sockdeo H., Turner K.J., Wood C.R.;
RT "Molecular cloning of murine FLT and FLT4.";
RL Oncogene 8:2293-2298(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94181281; PubMed=8134130;
RA Choi K., Wall C., Hanraty R., Keller G.;
RT "Isolation of a gene encoding a novel receptor tyrosine kinase from
RT differentiated embryonic stem cells.";
RL Oncogene 9:1261-1266(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Lung;
RX MEDLINE=98201626; PubMed=9524283;
RA Kondo K., Hiratauka S., Subbalakshmi E., Matsushige H., Shibuya M.;
RT "Genomic organization of the flt-1 gene encoding for vascular
RT endothelial growth factor (VEGF) receptor-1 suggests an intimate
RT evolutionary relationship between the 7-ig and the 5-ig tyrosine
RT kinase receptors.";
RL Gene 208:297-305(1998).
CC - FUNCTION: RECEPTOR FOR VEGF AND PLGF. HAS A TYROSINE-PROTEIN
CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF

```

```

CC VASCULAR PERMEABILITY.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases.
CC CSF-1/PDGF receptor subfamily.
CC -1- SIMILARITY: Contains 7 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; L07297; AAA40078.1; -
DR EMBL; X79568; CAA5311.1; -
DR EMBL; D88689; BAA24498.1; -
DR PIR; I78875; I78875.
DR PIR; S49010; S49010.
DR HSSP; P17948; IQSV.
DR MGD; MGI.95558; P1C1.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig C2.
DR InterPro: IPR001824; Receptor_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR008266; Tyr_kinase_AS.
DR Pfam; PF00047; Ig; 6.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00408; IGC2; 2.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS50835; IG LIKE; 5.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Angiogenesis; Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; Receptor; Transmembrane; Signal; Immunoglobulin domain;
KW Repeat; Glycoprotein.
FT CHAIN 1 22
FT SIGNAL 23 1333
FT DOMAIN 23 759
FT TRANSMEM 760 781
FT DOMAIN 782 1333
FT DOMAIN 32 124
FT DOMAIN 152 215
FT DOMAIN 231 328
FT DOMAIN 334 429
FT DOMAIN 430 550
FT DOMAIN 557 656
FT DOMAIN 662 748
FT DOMAIN 828 1158
FT NP_BIND 834 842
FT BINDING 862 862
FT ACT_SITE 1022 1022
FT MOD_RES 1053 1053
FT DISULFID 53 108
FT DISULFID 159 208
FT DISULFID 253 312
FT DISULFID 455 536
FT DISULFID 578 637
FT DISULFID 683 732
FT CARBOHYD 101 101
FT CARBOHYD 165 165
FT CARBOHYD 197 197
FT CARBOHYD 252 252
FT CARBOHYD 324 324
FT CARBOHYD 418 418
FT CARBOHYD 475 475

```

```

FT CARBOHYD 517 517 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 626 626 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 MISSING (IN REF. 2).
FT CONFLICT 211 211 MISSING (IN REF. 2).
FT CONFLICT 245 245 H -> L (IN REF. 2).
FT CONFLICT 603 603 H -> N (IN REF. 2).
FT CONFLICT 609 615 KMTTDD -> NHHSS (IN REF. 2).
FT CONFLICT 696 696 F -> L (IN REF. 2).
FT CONFLICT 734 734 A -> S (IN REF. 2).
FT CONFLICT 765 765 C -> Y (IN REF. 2).
FT CONFLICT 820 820 K -> N (IN REF. 2).
FT CONFLICT 1009 1009 G -> R (IN REF. 2).
FT CONFLICT 1181 1181 S -> G (IN REF. 3).
FT CONFLICT 1181 1181 S -> N (IN REF. 2).
FT CONFLICT 1193 1194 LP -> RG (IN REF. 2).
FT CONFLICT 1278 1279 KS -> PR (IN REF. 2).
SQ SEQUENCE 1333 AA; 149875 MM; C06533B7ECBC404C CRC64;

Query Match 7.4%; Score 97.5; DB 1; Length 1333;
Best Local Similarity 23.8%; Pred. No. 13;
Matches 45; Conservative 30; Mismatches 71; Indels 43; Gaps 9;

QY 33 KEDTVELCTGASQKSIQFHMKNNSNOIKILNGSGFLTGSPKLNDRADSRRLMPQGNF 92
Db 570 EGGDLKSLCVNPFYRDTW---ILKRTVNNTHMHSISKQKMTTQDYSTL----- 620
QY 93 PLIKNKLIKEDSDPTVYCEV-----EDQKEEYVLVFGLTNSDTHLQSGSLTLTLESP 146
Db 621 NLVTKNVSLSDGETVACRANITYTGEDIARKTEVLRDSEA---PILQNLSS---DYEVS 674
QY 147 PGSSPSVQCRSPRG-----KN---IQ-----GKTLVSQLELDQSGTWCT 185
Db 675 ISGSLTLDQA-RGVAPQITWFKNNHKIQDEPGIILGSGNSTLFLERVTEDEGVYRCR 733
QY 186 VLQNKQVE 194
Db 734 ATNKGAVE 742

RESULT 86
UN89 CAEBL STANDARD; PRT; 6632 AA.
AC 001761; Q17362;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
GN UNC-89 OR CO9D.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN=Briscot N2;
RX MEDLINE=96180278; PubMed=8603916;
RA Benian G.M., Tinley T.L., Tang X., Borodovsky M.;
RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line
RT assembly, encodes a giant modular protein composed of Ig and signal
RT transduction domains";
RL J. Cell Biol. 132:835-848(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Du Z., Le T.T., Nelson R.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

```


Db	3693	KETFAVAGDPTVLECKYKNKESHQPKFKNDQPEI----	CGHM-----	QLE	3735
Qy		86 LWDQGNFPLLIKNIKIEDSDTYICE-----	VED-----	QKEE	117
Db	3736	VLEEGNFKLTITQNKKEDEVGAYRCEAVNVACKANTMADLKIQFAKYEHEVTDSEGOLEE	3795		
Qy		118 V-QLLVGLGTANSPTHTLLOG-----	QSLLTLESPPGSSPSVOCR-----	SPRGKNI	163
Db	3796	IGQPEYTGDDTSSKTTDGRGAPPEVVELLRSCVTIEK----	QQAIIKCKYKKEPPKIKMT	3851	
Qy		164 QGKTLISV-----	QLELQDSGTWTC	185	
Db	3852	KEGKEVMSARVRAEHKDDGLTIT	3876		
RESULT 87					
FGRL MOUSE					
ID	FGRL MOUSE	STANDARD;	PRT;	822 AA.	
AC	P16092; Q01736; Q61562;				
DT	01-MAY-1990 (Rel. 14, Created)				
DT	01-MAY-1991 (Rel. 18, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
DE	Basic fibroblast growth factor receptor 1 precursor (EC 2.7.1.112)				
DE	(fgrR-1) (bfgf-R) (wfr).				
GN	FGFR1 OR FLG.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RC	MEDLINE=90160373; PubMed=1689490;				
RX	Reid H.H., Wilks A.F., Bernard O.;				
RT	"Two forms of the basic fibroblast growth factor receptor-like mRNA				
RT	are expressed in the developing mouse brain.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 87:1596-1600(1990).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BALB/c; TISSUE=Brain;				
RX	MEDLINE=90265603; PubMed=2161096;				
RA	Satrian A., Avioli A., Orr-Urtreger A., Neufeld G., Lonai P.,				
RA	Girol D., Yarden Y.;				
RT	"The murine flg gene encodes a receptor for fibroblast growth				
RT	factor.";				
RL	Oncogene 5:635-643(1990).				
RN	[3]				
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.				
RX	MEDLINE=91207411; PubMed=1708247;				
RA	Kouhara H., Kasayama S., Saito H., Matsumoto K., Sato B.;				
RT	"Expression cDNA cloning of fibroblast growth factor (FGF) receptor				
RT	in mouse breast cancer cells: a variant form in FGF-responsive				
RT	transformed cells.";				
RL	Biochem. Biophys. Res. Commun. 176:31-37(1991).				
RN	[4]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).				
RX	MEDLINE=90272715; PubMed=2161540;				
RA	Maniukhant A., Moscatelli D., Talarico D., Levytska V., Basilico C.;				
RT	"A murine fibroblast growth factor (FGF) receptor expressed in CHO				
RT	cells is activated by basic FGF and Kaposi FGF.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 87:4378-4382(1990).				
RN	[5]				
RP	SEQUENCE OF 1-15 FROM N.A., AND ALTERNATIVE SPLICING.				
RX	MEDLINE=95100926; PubMed=7802632;				
RA	Harada T., Saito H., Kouhara H., Kurebayashi S., Kasayama S.,				
RA	Terakawa N., Kishimoto T., Sato B.;				
RT	"Murine fibroblast growth factor receptor 1 gene generates multiple				
RT	messenger RNAs containing two open reading frames via alternative				
RT	splicing.";				
RL	Biochem. Biophys. Res. Commun. 205:1057-1063(1994).				
CC	-1- FUNCTION: Receptor for basic fibroblast growth factor. A shorter				
CC	form of the receptor could be a receptor for acidic FGF (aFGF).				
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein				

CC		tyrosine phosphate.
CC	-1-	SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1-	ALTERNATIVE PRODUCTS:
CC		Event=Alternative splicing; Named isoforms=3;
CC	Name=1;	
CC	IsoId=P16092-1;	Sequence=Displayed;
CC	Name=2;	
CC	IsoId=P16092-2;	Sequence=VSP_002962;
CC	Name=3;	Synonyms=Variant;
CC	IsoId=P16092-3;	Sequence=VSP_002961, VSP_002963;
CC	-1-	SIMILARITY: Belongs to the fibroblast growth factor receptor family.
CC	-1-	SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
CC		-----
DR	EMBL; M28988;	AAA37290.1; . -
DR	EMBL; X51893;	CAA36175.1; . -
DR	EMBL; M65053;	AAA37620.1; . -
DR	EMBL; M33760;	AAA37622.1; . -
DR	EMBL; S74765;	AAB32845.1; ALT_SEQ.
DR	PIR; A34849;	TWMSFG.
DR	HSSP; P13362;	IRGK.
DR	MGI; MGI:95542;	Fgf1.L
DR	GO; GO:00074420;	P:brain development; IMP.
DR	GO; GO:0042472;	P:inner ear morphogenesis; IMP.
DR	GO; GO:0007435;	P:salivary gland morphogenesis; IMP.
DR	InterPro; IPRO07110;	Ig-Like.
DR	InterPro; IPRO03598;	Ig_C2.
DR	InterPro; IPRO00719;	Prot_kinase.
DR	InterPro; IPRO01245;	Tyr_kinase.
DR	InterPro; IPRO08266;	Tyr_kinase_AS.
DR	Pfam; PF00069;	pkinase; 1.
DR	Pfam; PF00047;	Ig; 3.
DR	PRINTS; PRO0109;	TYRKINASE.
DR	ProDom; PD000001;	Prot_kinase; 1.
DR	SMART; SMO0408;	IGC2; 3.
DR	SMART; SMO0219;	TYKC; 1.
DR	PROSITE; PS50835;	Ig_LIKE; 3.
DR	PROSITE; PS01017;	PROTEIN KINASE_ATP; 1.
DR	PROSITE; PS50011;	PROTEIN KINASE_DOM; 1.
DR	PROSITE; PS01019;	PROTEIN KINASE_TYR; 1.
KW	Receptor; Glycoprotein;	Tyrosine-protein kinase; ATP-binding;
KW	Transferase; Phosphorylation;	Transmembrane; Signal;
KW	Immunoglobulin domain;	Repeat; Alternative splicing.
FT	SIGNAL	1 21
FT	CHAIN	22 822
FT		BASIC FIBROBLAST GROWTH FACTOR RECEPTOR
FT		1.
FT	DOMAIN	22 376
FT	TRANSMEM	377 397
FT		POTENTIAL.
FT	DOMAIN	398 822
FT		CYTOSOLASMIC (POTENTIAL).
FT	DOMAIN	25 119
FT		IG-LIKE C2-TYPE 1.
FT	DOMAIN	158 246
FT		IG-LIKE C2-TYPE 2.
FT	DOMAIN	255 357
FT		IG-LIKE C2-TYPE 3.
FT	DOMAIN	478 767
FT		PROTEIN KINASE.
FT	NP BIND	484 492
FT		ATP (BY SIMILARITY).
FT	BINDING	514 514
FT		ATP (BY SIMILARITY).
FT	ACT SITE	623 623
FT		BY SIMILARITY.
FT	MOD_RES	654 654
FT		PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	DISULFID	55 101
FT		POTENTIAL.
FT	DISULFID	178 230
FT		POTENTIAL.
FT	CARBOHYD	277 341
FT		POTENTIAL.
FT	CARBOHYD	77 77
FT		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	117 117
FT		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	227 227
FT		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	240 240
FT		N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	264 264
FT		N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match	Best Local Similarity	7.4%; Score 97; DB 1; Length 822;	
Matched 57; Conservative	32; Mismatches 92; Indels 106; Gaps 10;		
FT CARBOHYD	296	296	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD	317	317	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD	330	330	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT VANSPLIC	30	30	Q -> QGSSWPLVMAAA (In isoform 3).
FT VANSPLIC	31	119	Missing (In isoform 2).
FT VANSPLIC	148	149	Missing (In isoform 3).
FT VANSPLIC	229	229	FTID=VSP 002962.
FT VANSPLIC	256	258	FTID=VSP 002962.
FT VANSPLIC	270	270	FTID=VSP 002962.
FT VANSPLIC	387	387	FTID=VSP 002962.
FT VANSPLIC	440	440	FTID=VSP 002962.
FT VANSPLIC	508	508	FTID=VSP 002962.
FT VANSPLIC	544	544	FTID=VSP 002962.
FT VANSPLIC	756	756	FTID=VSP 002962.
FT VANSPLIC	765	765	FTID=VSP 002962.
FT VANSPLIC	822 AA;	91980 MM;	D5A4635FA680926B CRC64;
Query Match	Best Local Similarity	7.4%; Score 97; DB 1; Length 822;	
Matched 57; Conservative	32; Mismatches 92; Indels 106; Gaps 10;		
15 QALLPATQGNKVVGLKKGDVVELTCTASQKSIQIFNH-KNSNRIKILNGSGSFLTGCP	73		
164 KLHAPPAAL-----KTVKPKCPSSGTPNPPTLRMLKNGKFKPDHRIGYKV---	208		
74 SKLNRADRSRLW-----DQGNPLIIRKILKIEDSPYICEVEDQKEVQLVVF	123		
209 -----RATWIIIMDSVPSKQNYTCIVENYSGSINTYQLDVERSPHPILOA	259		
124 GLTANSDTHLLQGQSLTTLTLESPP-----GSS-----PSVQCRSPRGKNI-	163		
260 GLPANKTVALGSGNVFMCKVYSDPQPHIOWLKHIVNGSKIGIDNLPYQILKTAGVNTT	319		
164 -QGCGTTLVSQLELDGSGTWC-----TYLQNGKVEREKRIDVPASALP	207		
320 DKEMEVLLHRLRVSPFDAGEYTCITLGNISGLSHSAMLTVL-----	360		
208 APTGASALPDPQTASALPDPASALPALAVISFLGLGAGVCL	254		
361 -----ALEBRPAMTSPLEYLEIITICTGAFI-ISCML	391		
RESULT 88			
NCAL_CHICK	STANDARD;	PRT; 1091 AA.	
AC P13590; Q90918; Q90919;			
DT 01-JAN-1990 (Rel. 13, Created)			
DT 15-JUL-1998 (Rel. 36, Last sequence update)			
DT 10-OCT-2003 (Rel. 42, Last annotation update)			
DE Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM			
DE 180)			
OS Gallus gallus (Chicken).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC Gallus.			
OX NCBI_Taxid=9031;			
LN [1]			
RN SEQUENCE OF 1-175 FROM N.A., AND PARTIAL SEQUENCE.			
RP MEDLINE=87206190; PubMed=3576199;			
RA Cunningham B.A., Hemperly J.J., Murray B.A., Prediger E.A.,			
RA Brackebury R., Edelman G.M.;			
RT "Neural cell adhesion molecule: structure, immunoglobulin-like			
RT domain, cell surface modulation, and alternative RNA splicing.";			
RL Science 236:799-806(1987).			
RN [2]			
RN SEQUENCE OF 128-1091 FROM N.A., AND PARTIAL SEQUENCE.			
RP MEDLINE=86206089; PubMed=3458261;			
RA Hemperly J.J., Murray B.A., Edelman G.M., Cunningham B.A.;			
RT "Sequence of a cDNA clone encoding the polysialic acid-rich and			
RT cytoplasmic domains of the neural cell adhesion molecule N-CAM";			

```

RL  Proc. Natl. Acad. Sci. U.S.A. 83:3037-3041(1986).
RN  [3]
RP  SEQUENCE FROM N.A. (ISOFORM D.).
RA  MEDLINE=87092340; PubMed=3467341;
RT  Hemperly J.J., Edelman G.M., Cunningham B.A.;
RT  "cDNA clones of the neural cell adhesion molecule (N-CAM) lacking a
RT  membrane-spanning region consistent with evidence for membrane
RT  attachment via a phosphatidylinositol intermediate.";
RL  Proc. Natl. Acad. Sci. U.S.A. 83:9822-9826(1986).
RN  [4]
RP  SEQUENCE OF 810-1069 FROM N.A.
RA  MEDLINE=87033934; PubMed=3771645;
RT  Murray B.A., Owens G.C., Prediger E.A., Crossin K.L.,
RT  Cunningham B.A., Edelman G.M.;
RT  "Cell surface modulation of the neural cell adhesion molecule
RT  resulting from alternative mRNA splicing in a tissue-specific
RT  developmental sequence.";
RL  J. Cell Biol. 103:1431-1439(1986).
RN  [5]
RP  SEQUENCE OF 1-17 FROM N.A.
RA  MEDLINE=93122797; PubMed=1478668;
RT  Cowell G., Li B., Forrest D., Brackenbury R.;
RT  "Conserved regulatory elements in the promoter region of the N-CAM
RT  gene.";
RL  Genomics 14:875-882(1992).
RN  [6]
RP  SEQUENCE OF 1-17 FROM N.A.
RA  STRAIN=White leghorn; TISSUE=Erythrocyte;
RT  Sasser M., Covault J.;
RT  Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
RL  - FUNCTION: This protein is a cell adhesion molecule involved in
CC  neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC  neurites, etc.
CC  -1- ALTERNATIVE PRODUCTS:
CC  Event=Alternative splicing; Named isoforms=4;
CC  Name=A; Synonyms=N-CAM 180;
CC  IsoId=P13590-1; Sequence=Displayed;
CC  Name=B; Synonyms=N-CAM 140;
CC  IsoId=P13590-2; Sequence=VSP_002585;
CC  Name=C;
CC  IsoId=P13590-3; Sequence=VSP_002586;
CC  Name=D;
CC  IsoId=P13590-4; Sequence=VSP_002583, VSP_002584;
CC  -1- SIMILARITY: Contains 5 Immunoglobulin-like C2-type domains.
CC  -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC  -----
CC  This Swiss-Prot entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M15861; AAB59958.1; JOINED.
DR  EMBL; M15860; AAB59958.1; JOINED.
DR  EMBL; M15922; AAB59958.1; JOINED.
DR  EMBL; M15923; AAB59958.1; JOINED.
DR  EMBL; M15924; AAB59958.1; JOINED.
DR  EMBL; M21178; AAB59958.1; JOINED.
DR  EMBL; M21179; AAB59958.1; JOINED.
DR  EMBL; M21180; AAB59958.1; JOINED.
DR  EMBL; M15929; AAB59958.1; JOINED.
DR  EMBL; M15930; AAB59958.1; JOINED.
DR  EMBL; M15931; AAB59958.1; JOINED.
DR  EMBL; M15932; AAB59958.1; JOINED.
DR  EMBL; M15933; AAB59958.1; JOINED.
DR  EMBL; M15934; AAB59958.1; JOINED.
DR  EMBL; L29437; AAB59958.1; JOINED.
DR  EMBL; M15935; AAB59958.1; JOINED.
DR  EMBL; M15937; AAB59958.1; JOINED.
DR  EMBL; M15938; AAB59958.1; JOINED.
DR  EMBL; M15939; AAB59958.1; JOINED.
DR  EMBL; M15939; AAB59958.1; JOINED.

```

FT	CARBOHYD	478	478	(POTENTIAL).
FT	VARSPLIC	702	726	STPTSGGLGAAVGLIVLIFVLL -> TLGSPSTSSFV
FT				SILLSVTLICIL (in isoform D).
FT	VARSPLIC	727	1091	/Frid=VSP 002583.
FT				Missing (in isoform D).
FT	VARSPLIC	771	809	/Frid=VSP_002584.
FT				Missing (in isoform B).
FT	VARSPLIC	810	1070	/Frid=VSP 002585.
FT				Missing (in isoform C).
SO	SEQUENCE	1091 AA,	117415 MW;	B151367002DD80BD CXC64;
 Query Match 7.4%; Score 97; DB 1; Length 1091; Best Local Similarity 21.4%; Pred. No. 11; Matches 51; Conservative 27; Mismatches 84; Indels 76; Gaps 10;				
OY	35 DTVELCTTASQKSIQHFMKN-----SNCKIKINGOGSFLTKGPSKLNDRADRSRLMDQ	89		
Dd	323 DQTITLEASGDPIPSITWTTSIRNINSNEKITL--DRIVRSHARVS-----	369		
OY	90 GNFPILIKNLKIDSDTYICEVEDOKEBVQLLVFGLTANSDTHL-----LOGQSL	139		
Dd	370 ----LTLKEIQTVDAGEYVCTAAN-----TIQQDSQAWYLEVAPKLGCPVA	413		
OY	140 TLTLESPPGSSPVOCR-----SPRKNIQGKGTLISVSOLE----	176		
Dd	414 VYTWE---GNQVAVITEVFAYPSAVISWFRDGQLLPBSNYSNKIKINTPASLYLEVTPDS	470		
OY	177 -ODSGTWTCVLQ--NOKKVEFKIDIVPRASALPAPEPGSALPDPOFASALPDPFAAS	231		
Dd	471 ENDFGNNTCAVMRIQESSEF---ILVQADPTSPSIDRVEEYSSARAVEPFDEPAT	525		
 RESULT 89 DCC HUMAN ID_DCC_HUMAN STANDARD; PRT; 1447 AA. P4146; Dt 01-NOV-1995 (Rel. 32, Created) Dt 01-NOV-1995 (Rel. 32, Last sequence update) Dt 10-OCT-2003 (Rel. 42, Last annotation update) De Tumor suppressor protein DCC precursor (Colorectal cancer suppressor). DCC. OS Homo sapiens (Human). OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eueleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. NCBI_Taxid=9606; CX [1] RN SEQUENCE FROM N.A. RP MEDLINE=95011532; PubMed=7926722; RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W., Vogelstein B.; RT "The DCC gene product in cellular differentiation and colorectal tumorigenesis." 1174-1183 (1994). RL Genes Dev. 8:1174-1183 (1994). RN [2] RP SEQUENCE OF 1-750 FROM N.A. RX MEDLINE=90100559; PubMed=2294591; RA Fearon E.R., Cho K.R., Ngro J.M., Kern S.E., Simons J.W., Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W., Vogelstein B.; RT "Identification of a chromosome 18q gene that is altered in colorectal cancers." Science 247:49-56(1990). RL [3] RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBLED EXONS). RX MEDLINE=9112517; PubMed=1991322; RA Niigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M., Oliner J.D., Kinzler K.W., Vogelstein B.; RT "Scrambled exons." RL Cell 64:607-613 (1991). RN [4] RP GENE STRUCTURE, AND VARIANT CARCINOMA HIS-1375. RX MEDLINE=94245241; PubMed=8186295;				

RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,
 RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;
 RT "The DCC gene: structural analysis and mutations in colorectal
 RT carcinomas";
 RN Genomics 19:525-531(1994).
 RN (5)
 RP VARIANT CARCINOMA THR-168, AND VARIANT GLY-201.
 RX MEDLINE=94243823; PubMed=8187090;
 RA Miyake S., Nagai K., Yoshino K., Ota M., Endo M., Yuasa Y.;
 RT "point mutations and allelic deletion of tumor suppressor gene DCC in
 RT human esophageal squamous cell carcinomas and their relation to
 RT metastasis";
 RL Cancer Res. 54:3007-3010(1994).
 CC -1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: FOUND IN AXONS OF THE CENTRAL AND PERIPHERAL
 CC NERVOUS SYSTEM AND IN DIFFERENTIATED CELL TYPES OF THE INTESTINE.
 CC -1- DISEASE: Colorectal tumors that lost their capacity to
 CC differentiate into mucus producing cells uniformly lack DCC
 CC expression. Inactivation of DCC due to allelic deletion and/or
 CC point mutations may cause both lymphatic and hematogenous
 CC metastasis of esophageal squamous cell carcinomas.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. DCC family.
 CC -1- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.
 CC -1- SIMILARITY: Contains 6 fibronectin type III domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X76132; CAAS3735.1; -;
 DR EMBL; M32292; AAA52751.1; -;
 DR EMBL; M32286; AAA52174.1; -;
 DR EMBL; M32288; AAA52175.1; ALT_SEQ.
 DR EMBL; M32290; AAA52176.1; -;
 DR EMBL; M63696; AAA52177.1; -;
 DR EMBL; M63700; AAA52178.1; -;
 DR EMBL; M63702; AAA52179.1; -;
 DR EMBL; M63718; AAA52180.1; -;
 DR EMBL; M63698; AAA52181.1; -;
 DR PIR; A54100; A54100.
 DR HSSP; P56276; ITLK.
 DR Genew; HGNC:2701; DCC.
 DR MIM; 120470; -;
 DR GO; GO:0004888; F:transmembrane receptor activity; TMS.
 DR GO; GO:0007409; P:axonogenesis; TMS.
 DR GO; GO:0006917; P:induction of apoptosis; TMS.
 DR InterPro; IPR008957; FN_III-like.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003962; FNIII subd.
 DR InterPro; IPR007110; IG-1-like.
 DR InterPro; IPR003598; IG_C2.
 DR Pfam; PF00041; fn3; 6.
 DR Pfam; PF00047; Ig; 4.
 DR PRINTS; PR00014; FNTYPEIII.
 DR SMART; SMO0060; FN3; 6.
 DR SMART; SMO0408; IGC2; 3.
 DR PROSITE; PS00835; IG-LIKE; 4.
 KW Glycoprotein; Immunoglobulin domain; Transmembrane; Signal;
 KW Repeat; Anti-oncogene; Disease mutation; Polymorphism.
 FT SIGNAL 1 25
 FT CHAIN 26 1447
 FT DOMAIN 26 1097 TUMOR SUPPRESSOR PROTEIN DCC.
 FT TRANSMEM 1098 1122 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 1123 1447 POTENTIAL.
 FT DOMAIN 26 135 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 139 229 IG-LIKE C2-TYPE 1.
 FT DOMAIN 234 326 IG-LIKE C2-TYPE 2.
 FT DOMAIN 331 416 IG-LIKE C2-TYPE 3.
 FT IG-LIKE C2-TYPE 4.

FT DOMAIN 426 522 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 525 618 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 619 716 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 722 816 FIBRONECTIN TYPE-III 4.
 FT DOMAIN 840 940 FIBRONECTIN TYPE-III 5.
 FT DOMAIN 941 1042 FIBRONECTIN TYPE-III 6.
 FT DISULFID 61 117 BY SIMILARITY.
 FT DISULFID 161 212 BY SIMILARITY.
 FT DISULFID 261 310 BY SIMILARITY.
 FT DISULFID 352 400 BY SIMILARITY.
 FT CARBOHYD 94 94 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 702 702 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 168 168 M->T (in esophageal carcinoma).
 FT VARIANT 168 168 /FtId=VAR_003909.
 FT VARIANT 201 201 R->G.
 FT VARIANT 201 201 /FtId=VAR_003910.
 FT VARIANT 1375 1375 P->H (in a colorectal carcinoma).
 FT CONFLICT 138 138 /FtId=VAR_003911.
 FT CONFLICT 233 329 MISSING (IN REF. 3).
 FT CONFLICT 421 421 MISSING (IN REF. 3).
 SQ SEQUENCE 1447 AA; 158456 MW; 4A8612766ED0471F CRC64;
 Query Match 7.4%; Score 97; DB 1; Length 1447;
 Best Local Similarity 23.7%; Pred. No. 15;
 Matches 45; Conservative 26; Mismatches 77; Indels 42; Gaps 8;
 Oy 34 GDTVELCTASQKRSIQFHMKNSQIKILNQGSFLRKPSKUNDRADSRRLSDQGNFP 93
 Db 154 GDTVLLKCEVIGSEMPFIHQKNOQ-----DLPIIP-----GDSRVVLPSC--A 196
 Oy 94 LIINKLKIEDSDTYICEVED-----QKEEVQL-----VFGLTANSDFTHLGOS 138
 Db 197 LQISRRLPGDIGIRCARNPASRTGNEAVRLSPGLHRLQYLFGRPSNVVAIRGKD 256
 Oy 139 LTL---TLSEPPGS-----SPYQCRSPKGNKIQGSKTVLSQLEQDSGTWTCTYKQN 189
 Db 257 AVLECCVSGYPPFPFTWLRGSEVQLRSKK-YSLDGSNLLISVTTDDSGMYTCVITYK 315
 Oy 190 QKTYEKKIDI 199
 Db 316 NENISASAEI 325
 RESULT 90
 KY1J HUMAN
 ID KY1J HUMAN STANDARD; PRT; 117 AA.
 AC P01602;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE IG kappa chain V-I region HK102 precursor (Fragment).
 GN IGKV1-5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81098966; PubMed=679204;
 RA Bentley D.L., Rabbits T.H.;
 RT "Human immunoglobulin variable region genes -- DNA sequences of two V
 RT kappa genes and a pseudogene";
 RL Nature 288:730-733(1980).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ibd-sib.ch/announce/
CC or send an email to license@ib-sib.ch).
CC -----
CC DR EMBL; J00245; AAA59087.1; -.
CC DR EMBL; Z00001; CAA77292.1; -.
CC DR PIR; A01882; K1HUI2.
CC DR HSSP; P01607; IREI.
CC DR Genew; HGNC:5741; IGKV1-5.
CC DR GO; GO:0005576; C:extracellular; NAS.
CC DR GO; GO:0003823; F:antigen binding; NAS.
CC DR GO; GO:0006955; P:immune response; NAS.
CC DR InterPro; IPR007110; IG-1 like.
CC DR InterPro; IPR003596; IG-V.
CC DR Pfam; PF00047; Ig; 1.
CC DR SMART; SM00406; IGV; 1.
CC DR PROSITE; PS50835; IG_LIKE; 1.
CC DR K1 Immunoglobulin V region; Signal.
CC FT CHAIN 1 22
CC FT DOMAIN 23 >117 IG KAPPA CHAIN V-I REGION HK102.
CC FT DOMAIN 23 45 FRAMEWORK-1.
CC FT DOMAIN 46 56 COMPLEMENTARITY-DETERMINING-1.
CC FT DOMAIN 57 71 FRAMEWORK-2.
CC FT DOMAIN 72 78 COMPLEMENTARITY-DETERMINING-2.
CC FT DOMAIN 79 110 FRAMEWORK-3.
CC FT DOMAIN 111 >117 COMPLEMENTARITY-DETERMINING-3.
CC FT DISULFID 45 110 BY SIMILARITY.
CC FT NON TER 117
CC SQ SEQUENCE 117 AA; 12768 MW; ADDPFA40NF1A49B CRC64;
CC
CC Query Match 7.3%; Score 96.5; DB 1; Length 117;
CC Best Local Similarity 31.7%; Pred. No. 0.77;
CC Matches 38; Conservative 11; Mismatches 52; Indels 19; Gaps 5;
CC
CC Oy 1 MNRGVPFHLILYLOLALLPPA-----TQGNKVYLGKKDPTVELCTASOKSQTQFM- 53
CC Db 1 MDMRPAC-LLGLLLMLPQAKCIQMTQSPSTLSASVDGRVYTTCASQSSISWLAWY 58
CC
CC Oy 54 --KNSNQIKIIGNQGSFLTKG-PSKLNDRADSRSLMDQGNFLLIKNLKIEDSPYICE 110
CC Db 59 QOKPGAKKLILVDASLESIGVPSRFSGSGT-----EFTLTISLQPDDEFATYVQC 111
CC
CC RESULT 91
CC PIGR_RAT ID STANDARD; PRT; 769 AA.
CC AC P15083;
CC DT 01-APR-1990 (Rel. 14, Created)
CC DT 01-APR-1990 (Rel. 14, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Polymeic-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)
CC [Contains: Secretory component].
CC GN PIGR.
CC OS Rattus norvegicus (Rat).
CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC CX NCBI_Txid=10116;
CC RN (1)
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Liver;
CC RX MEDLINE=89378226; PubMed=2776882;
CC RA Bartling G., Brake B., Braghetta P., Luzio J.P., Stanley K.K.;
CC RT "Intracellular targeting signals of polymeic immunoglobulin
CC RT receptors are highly conserved between species.";
CC RL FEBS Lett. 254:177-183 (1989).
CC CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APOICAL SURFACE.
CC CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC CC TRANSMEMBRANE SEGMENT.
CC CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Also secreted.
CC CC -!- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.

```

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X15741; CAA33758.1; -.
DR PIR: S05407; QRR7GS.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003599; Ig.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00409; Ig; 5.
DR PROSITE: PSS0835; IG_LIKE; 2.
KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 19 769 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CDS 19 611 SECRETORY (POTENTIAL).
FT DOMAIN 19 643 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 644 666 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 667 769 POTENTIAL.
FT DOMAIN 21 126 IG-LIKE V-TYPE 1.
FT DOMAIN 135 237 IG-LIKE V-TYPE 2.
FT DOMAIN 240 341 IG-LIKE V-TYPE 3.
FT DOMAIN 353 457 IG-LIKE V-TYPE 4.
FT DOMAIN 463 563 IG-LIKE V-TYPE 5.
FT DISULFID 40 110 POTENTIAL.
FT DISULFID 152 220 POTENTIAL.
FT DISULFID 257 324 POTENTIAL.
FT DISULFID 370 440 POTENTIAL.
FT DISULFID 484 546 POTENTIAL.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 769 AA; 84798 MW; 5F849303400255A7 CRC64;

Query Match 7.3%; Score 96.5; DB 1; Length 769;
Best Local Similarity 21.9%; Pred. No. 7.8;
Matches 59; Conservative 46; Mismatches 107; Indels 57; Gaps 13;

11 LVVLQLALLPAATGKNVLTGKGGDYELTCTASQKSIQF---HK---KNSNQIKLG 63
343 LFVNEESTIP---NSRSVVGVTGSGVAIVCPYNPKESSSLKTKWCHWADNGRCPYLVG 399
64 NQGSFLLTKGFSKINDRADSRRLMDQ---GNPFLINKLIEDSYTICEVEDQ---KEE 117
400 TQ-----ALVGEVGEGRLLALFDQGSAGAYIILNQLTQDGSFGWCLTDGSRMRTT 451
118 VQLVLPGLTANSDTHL-----LQGSGLTLLTESPPG--SSPSVQR-SPRGNI----- 163
452 IELQVAVATKPKDELVTLPONATVIGTFTFISCHYPCPKFYQSEKYKWKMSDGHILPSH 511
164 -QGKTLVS-----QLTEQDSGWTCTYLAQOKKVEFKIDIVPRASALPAP 209
512 DEGRKQSSVSCDSQSIVSMTLNPKVEDGEMTWCGYKEGVYGETTAIVY---AYVER 567
210 PTGSALPDPOFASA-LPDPPAASALPAL 237
568 TRGSPHINPTDANARAKADABEEAMESSV 596

RESULT 92
SN MOUSE STANDARD; PRT; 1694 AA.
AC Q62230; O55216; Q62228; Q62229;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Stialodhesin precursor (Stialic acid binding Ig-like lectin-1) (Siglec-
1) (Sheep erythrocyte receptor) (SER).

```

GN SN OR SA.
 OS Mus musculus (Mouse).
 OC Euteleostomi; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=C57BL/6; TISSUE=Macrophage;
 RX MEDLINE=95009950; PubMed=7925291;
 RA Crocker P.R., Mucklow S., Bouksoo V., McWilliam A., Willis A.C.,
 RA Gordon S., Milon G., Kelm S., Bradfield P.,
 RT "Sialoadhesin, a macrophage sialic acid binding receptor for
 RT haemopoietic cells with 17 immunoglobulin-like domains.",
 RL EMO J. 13:4490-4503 (1994).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96051930; PubMed=9383289;
 RA Mucklow S., Gordon S., Crocker P.R.,
 RT "Characterization of the mouse sialoadhesin gene, Sn.",
 RT Mamm. Genome 8:934-937 (1997).
 RL [3]
 RN CHARACTERIZATION
 RC STRAIN=C57BL/6; TISSUE=spleen;
 RX MEDLINE=91266893; PubMed=2050106;
 RA Crocker P.R.,
 RT "Purification and properties of sialoadhesin, a sialic acid-binding
 RT receptor of murine tissue macrophages.",
 RL EMO J. 10:1661-1669 (1991).
 RN [4]
 RP SIALIC ACID BINDING.
 RX MEDLINE=95179521; PubMed=7533044;
 RA Kelm S., Pelz A., Schauer R., Filbin M.T., Tang S., de Bellard M.E.,
 RA Schnaar R.L., Mahoney J.A., Hartnell A., Bradfield P., Crocker P.R.,
 RT "Sialoadhesin, myelin-associated glycoprotein and CD22 define a new
 RT family of sialic acid-dependent adhesion molecules of the
 RT immunoglobulin superfamily.",
 RL Curr. Biol. 4:965-972 (1994).
 RN [5]
 RP BINDING TO SPN.
 RX MEDLINE=21136329; PubMed=11285599;
 RA van den Berg T.K., Nath D., Zillner H.J., Vestweber D., Fukuda M.,
 RA Van Die I., Crocker P.R.,
 RT "CD43 functions as a T cell counterreceptor for the macrophage
 RT adhesion receptor sialoadhesin (Siglec-1).",
 RL J. Immunol. 166:3637-3640 (2001).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 20-138.
 RX MEDLINE=96325385; PubMed=9660955;
 RA May A.P., Robinson R.C., Vinson M., Crocker P.R., Jones E.Y.,
 RT "Crystal structure of the N-terminal domain of sialoadhesin in
 RT complex with 3' sialyllactose at 1.85 A resolution.",
 RL Mol. Cell 1:719-728 (1998).
 RN [7]
 RP STRUCTURE OF 20-138 BY NMR, AND MUTAGENESIS OF TRP-21 AND ARG-116.
 RX MEDLINE=99321481; PubMed=10393093;
 RA Crocker P.R., Vinson M., Kelm S., Drickamer K.,
 RT "Molecular analysis of sialoside binding to sialoadhesin by NMR and
 RT site-directed mutagenesis.",
 RL Biochem. J. 341:355-361 (1999).
 CC -1- FUNCTION: Macrophage-restricted adhesion molecule that mediates
 CC sialic-acid dependent binding to lymphocytes, including
 CC granulocytes, monocytes, natural killer cells, B-cells and CD8 T-
 CC cells (By similarity). Preferentially binds to alpha2,3-linked
 CC sialic acid. Binds to SPN/CD43 on T-cells. May play a role in
 CC hemopoiesis.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and
 CC soluble (isoforms 2 and 3).
 CC -1- ALTERNATIVE PRODUCTS: Named isoforms=3;
 CC Event=Alternative splicing;
 CC Name=1;
 CC IsoId=Q62230-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q62230-2; Sequence=VSP_002573, VSP_002574;

CC Name=3;
 CC IsoId=Q62230-3; Sequence=VSP_002575, VSP_002576;
 CC -1- TISSUE SPECIFICITY: Expressed by macrophages in various tissues.
 CC Highest expression in spleen and lymph node with lower amounts in
 CC lung, liver, bone marrow, heart and skin. No expression in thymus,
 CC kidney, brain or small intestine.
 CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
 CC (sialic acid binding Ig-like lectin) family.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z36293; CAAB5290.1; -;
 DR EMBL; Z36233; CAAB5268.1; -;
 DR EMBL; Z36234; CAAB5269.1; -;
 DR EMBL; U92842; AAB95641.1; -;
 DR EMBL; U92833; AAB95641.1; JOINED.
 DR EMBL; U92834; AAB95641.1; JOINED.
 DR EMBL; U92835; AAB95641.1; JOINED.
 DR EMBL; U92837; AAB95641.1; JOINED.
 DR EMBL; U92838; AAB95641.1; JOINED.
 DR EMBL; U92839; AAB95641.1; JOINED.
 DR EMBL; U92840; AAB95641.1; JOINED.
 DR EMBL; U92841; AAB95641.1; JOINED.
 DR PIR; S50065; S50065.
 DR PDB; 1QFC; 16-APR-99.
 DR PDB; 1QFP; 16-APR-99.
 DR MGD; MGI:99668; Sn.
 DR GO; GO:0016021; C:integral to membrane; ISS.
 DR GO; GO:0005529; F:sugar binding; ISS.
 DR GO; GO:0016337; P:cell-cell adhesion; ISS.
 DR GO; GO:0007160; P:cell-matrix adhesion; ISS.
 DR GO; GO:0006954; P:inflammatory response; ISS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00047; Ig; 16.
 DR SMART; SM00408; IGc2; 5.
 DR PROSITE; PS50835; IG-LIKE; 14.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
 KW Immunoglobulin domain; Repeat; Alternative splicing; 3D-structure.
 FT SIGNAL 1 19
 FT CHAIN 20 1694
 FT DOMAIN 20 1638
 FT TRANSMEM 1639 1659
 FT DOMAIN 1660 1694
 FT DOMAIN 20 136
 FT DOMAIN 153 235
 FT DOMAIN 239 321
 FT DOMAIN 326 406
 FT DOMAIN 416 508
 FT DOMAIN 509 594
 FT DOMAIN 602 701
 FT DOMAIN 704 781
 FT DOMAIN 795 890
 FT DOMAIN 894 972
 FT DOMAIN 979 1078
 FT DOMAIN 1080 1160
 FT DOMAIN 1171 1263
 FT DOMAIN 1244 1336
 FT DOMAIN 1341 1438
 FT DOMAIN 1441 1519
 FT DOMAIN 1533 1626
 FT SITE 827 829
 FT DISULFID 36 166
 SILOADHESIN.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 IG-LIKE V-TYPE.
 IG-LIKE C2-TYPE 1.
 IG-LIKE C2-TYPE 2.
 IG-LIKE C2-TYPE 3.
 IG-LIKE C2-TYPE 4.
 IG-LIKE C2-TYPE 5.
 IG-LIKE C2-TYPE 6.
 IG-LIKE C2-TYPE 7.
 IG-LIKE C2-TYPE 8.
 IG-LIKE C2-TYPE 9.
 IG-LIKE C2-TYPE 10.
 IG-LIKE C2-TYPE 11.
 IG-LIKE C2-TYPE 12.
 IG-LIKE C2-TYPE 13.
 IG-LIKE C2-TYPE 14.
 IG-LIKE C2-TYPE 15.
 IG-LIKE C2-TYPE 16.
 CELL ATTACHMENT SITE (POTENTIAL).
 BY SIMILARITY.

```

FT DISULFID 41 98 BY SIMILARITY.
FT DISULFID 160 218 BY SIMILARITY.
FT DISULFID 263 306 BY SIMILARITY.
FT DISULFID 347 391 BY SIMILARITY.
FT DISULFID 434 492 BY SIMILARITY.
FT DISULFID 532 576 BY SIMILARITY.
FT DISULFID 625 685 BY SIMILARITY.
FT DISULFID 725 770 BY SIMILARITY.
FT DISULFID 813 872 BY SIMILARITY.
FT DISULFID 911 955 BY SIMILARITY.
FT DISULFID 1000 1062 BY SIMILARITY.
FT DISULFID 1102 1144 BY SIMILARITY.
FT DISULFID 1188 1236 BY SIMILARITY.
FT DISULFID 1276 1319 BY SIMILARITY.
FT DISULFID 1362 1421 BY SIMILARITY.
FT DISULFID 1462 1508 BY SIMILARITY.
FT DISULFID 1551 1610 BY SIMILARITY.
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 583 583 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 693 693 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 722 722 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 882 882 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1089 1089 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1099 1099 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1246 1246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1459 1459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1473 1473 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 326 340 MAEYKMPAGFVLN -> SSWMLKGPVSGKH (in isoform 2).
FT VARSPLIC 341 1694 /Frid=VSP 002573.
FT VARSPLIC 1528 1598 Missing (in isoform 2).
FT VARSPLIC 1598 /Frid=VSP 002574.
FT VARSPLIC 1599 1694 YPPTPLTIVFEQGHGGLDGRUNSEPLALTLRGSQ
FT MUTAGEN 21 21 LVASNQJHDAVTKPHINIVAPNALRVIE -> CEYEPIS
FT MUTAGEN 116 116 ALCLSLHGTGYQAFSSAOSGKFGKGLRTLASLACWMP
FT MUTAGEN 116 116 VSMGLPALKWRILPLPMDEYR (in isoform 3).
FT VARSPLIC 1599 1694 /Frid=VSP 002575.
FT VARSPLIC 1599 1694 Missing (in isoform 3).
FT MUTAGEN 21 21 W->: LOSS OF SIMILIC ACID BINDING.
FT MUTAGEN 116 116 R->A: LOSS OF SIMILIC ACID BINDING.
FT MUTAGEN 116 116 R->L: 10-FOLD LOSS IN AFFINITY TO SIMILIC
Query Match 7.3% Score 96.5; DB 1; Length 1694;
Best Local Similarity 24.8%; Pred. No. 20;
Matches 75; Conservative 37; Mismatches 106; Indels 85; Gaps 17;

```

```

Db 284 LGV 286
LU_HUMAN
ID LU_HUMAN STANDARD; PRT; 628 AA.
AC P50895;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Lutheran blood group glycoprotein precursor (B-CAM cell surface glycoprotein) (Auberger B antigen) (Fg/G253 antigen).
GN LU OR BCAM OR MSX19.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
OX [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-67 AND 182-203.
RC TISSUE=Placenta;
RX MEDLINE=95296337; PubMed=777537;
RA Parsons S.F., Mallinson G., Holmes C.H., Houlihan J.M., Simpson K.L., Mawby W.J., Spurr N.K., Warne D., Barclay A.N., Anstee D.J.;
RT "The Lutheran blood group glycoprotein, another member of the immunoglobulin superfamily, is widely expressed in human tissues and is developmentally regulated in human liver."
RL Proc. Natl. Acad. Sci. U.S.A. 92:5496-5500(1995).
RN [2]
RX SEQUENCE OF 1-588 FROM N.A.
RX MEDLINE=95042297; PubMed=7954395;
RA Campbell I.G., Foulkes W.D., Senger G., Trowdale J.,
RT "Molecular cloning of the B-CAM cell surface glycoprotein of epithelial cancers: a novel member of the immunoglobulin superfamily."
RL Cancer Res. 54:5761-5765(1994).
RN [3]
RX CARBOHYDRATE-LINKAGE SITE ASN-439.
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-J., Martin D.B., Abersold R.;
RT "Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry." Nat. Biotechnol. 21:660-666(2003).
RL Nat. Biotechnol. 21:660-666(2003).
CC -1- FUNCTION: PROBABLE RECEPTOR. MAY MEDIATE INTRACELLULAR SIGNALING.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: WIDE TISSUE DISTRIBUTION (HIGHEST IN THE PANCREAS AND VERY LOW IN BRAIN). CLOSELY ASSOCIATED WITH THE BASAL LAYER OF CELLS IN EPITHELIA AND THE ENDOTHELIUM OF BLOOD VESSEL WALLS.
CC -1- DEVELOPMENTAL STAGE: Is under developmental control in liver and may also be regulated during differentiation in other tissues.
CC -1- UPREGULATED following malignant transformation in other cell types.
CC -1- POLYMORPHISM: LU is responsible for the Lutheran blood group system.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.
CC -1- DATABASE: NAME=Blood group antigen mutation database;
CC NOTE=Lutheran (Lu) blood group system;
CC WWW="http://www.bdoc.aecom.yu.edu/bgmutc/lutheran.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC EMBL; X83425; CAA58449.1; -
CC DR EMBL; X80026; CAA56327.1; -
CC DR PIR; I37202; I37202.
CC DR PIR; I38000; I38000.

```